

TCAAAACGTT AGAGAAAGTT CTTCAAAAGT GCAGCCACAGA GCTGCTGGGC CCACCTGGCCG 3000
 TCCTGCATTT CTGGTTTCCA GACCCCAATG CCTCCCATTC GGATGGATCT CTGCGTTTTT 3060
 ATACGAGTGT TGCTAGGTT GCCCTTAT TTTATTTTC CCGTTGCGT TGCTATAGAT 3120
 GAAGGCTGAG GACAATCGTG TATATGTACT AGAATTTTT TATTAAGAA A

A145 Protein sequence:

Gene name: Cadherin 3, P-cadherin (placental)
 Unigene number: Hs.2877
 Probeset Accession #: X63629
 Protein Accession #: CAA45177
 Signal sequence: 1-24
 Transmembrane domain: 659-675

Cellular localization: plasma membrane

1 11 21 31 41 51
 MGLPRGFLAS LLLQVCWLIQ CAASEPCRAV FREAEVTLEA GGAQEPEGQA LGKVFMCPCG 60
 QEPALFSTDN DDFTVKNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120
 KGPFQRLNQ LKMKDRDTK IFYSITGPGA DSEPPGVFAV EKETGWLLEN KPLDRREBIK 180
 YELFGHAYSE NGASVEDEMN ISIIITDQND HKPKFTQDTF RGSVLEGVLP GTSVMQVTAT 240
 DEDDAIYFTN GVAVYSIHSQ EPKDPHDLMF TIHRSTGTLS VISSGLDREK VPEYTLTIQA 300
 TDMDGSGSTT TAVAVVEILD ANDNAPMEDP QKYRAHVPEV AVGHSEVQLT VTDLDAFNPS 360
 AWRATYLMG GDDGDHFTIT THPESNQGIL TTRKGLDPEA KNOHTLYVEV TNEAPFVLKL 420
 PTSTATIVVH VEDVNEAPVF VPPSKVVEVQ EGITPGEVPC VYTAEDPDKE NQKISYRILR 480
 DPAGLWAMSE DSGQVTAVGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLLTLLD 540
 VNDHGVPEBP RQITICNPSQ VRHVLNIDK DLSPHTSPFQ AQLTDDSDIY WTAEVNEEGD 600
 TVVLGLKKFL KQDYDVLELS LSHGKKEQL TVIRATVDCD HGHVETCPGP WKGGFLEVL 660
 GAVLALLFL LVLILLVRKK RKIKEPILLP EDDTRDNVYF YGEGGGGEEB QDYDITQLHR 720
 GLEARPEVVL RNDVAPTILF TPMYRFRFAN FDEIGNFIE NLKAANTDET APPYDTLLVF 780
 DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN EWSGRFKKLA DMYGGGDD

A146 DNA SEQUENCE:

Gene name: TTK protein kinase
 Unigene number: Hs.169840
 Probeset Accession #: M86699
 Nucleic Acid Accession #: NM_003318
 Coding sequence: 1026-3551 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 GGAATTCCTT TTTTITTTT TTTGAGATGG AGTTTCACTC TTGTTGGCCA GGCTGGAGTG 60
 CAATGGCACA ATCTCAGCTT ACTGCAACCT CGGCTCCCGG GGTTCAGAGG ATTCTCTGTC 120
 CTGAGCCTCT CAGTAGAGTG GGATTACAGG CATGTGCCAC CACCCCTGGC TAACCTAATT 180
 CTTTCTATT TATAGAGAT GGGGTTTCAC CATGTTGGTC AGGCTGGTCT TGAACCTCTG 240
 ACCTCAGGTG ATCCACTTGC CTGCGCCTCC CAAAGTGCTA GGATTACAGC CGTGAAACTG 300
 TGCCCGGCTG ATCTTTTTT TGTGTTGGA TTTTGAAC AGGCTCTCCC TTGGTCGCC 360
 AGGCTGGAGT GCAGTGGTGC GATCTTGGCT CACTATAACC TCCACCTCCT GGTTCAGT 420
 GATCTGCCA CTTCAGCTTC CTGAGTAGCT GTGATTACAG GCTGCAACA CCACACCCGG 480
 CTAATTTTG TATTTTATT AGAGACAGGG TTTCAACATG TTGGCCAGCG TGTCTCTCAA 540
 CTCCTGGACT CAGGGATCC GCCTGCCCTC ACTTCCCAAA GTCCCGAGAT TACAGGTGTG 600
 AGTCACCATG CTCGACCTTA TAATCTTAA GTCATTTTT CTGGTCCATT TCTTCTTAG 660
 GGYCTCACA ACAATCTGC ATTAGGCGGT ACAATAATCC TTAACCTCAT GATTACAAA 720
 AGGAAGATGA AGTGATTCTT GATTTAGAAA GGGGAAGTAG TAAGCCCACT GCACACTCCT 780
 GGATGATGAT CCTAAATCCA GATACAGTAA AAATGGGTTA TGGGAAGGTA GAATACAAA 840
 TTTGGTTTAA ATTAATATAT TAAATATCTA AAAACATTTT TGATATACAT GTTGATGTGA 900
 ATGTAAGACT GTACAGACTT OCTAGAAAAC AGTTTGGGTT CCATCTTTC ATTCCCCG 960
 TGCAGTTTTT TGTAGAAATG GAATCCAGG ATTTAAGTGG CAGAGATTTG ACAATGTATT 1020
 CCATAATGAA CAAAGTGAGA GACATTAAAA ATAAATTTAA AAATGAAGAC CTACTGATG 1080
 AACTAAGCTT CAAATAAATT TCTGCTGATA CTACAGATAA CTCGGGAAC GTTAACCAA 1140
 TTATGATGAT GGCACCAAC CCAGAGGACT GGTGAGTTT GTTGCTCAA CTAGAGAAA 1200
 ACAGTGTTCC GCTAAGTGAT GCTCTTTTAA ATAAATTGAT TGGTGGTAC AGTCAGCAA 1260
 TTGAAGCGCT TCCCCAGAT AATATGGCC AAAATGAGAG TTTTGCTAGA ATTCAAGTGA 1320
 GATTGCTGA ATTAAGAGCT ATTCAAGAGC CAGATGATGC ACGTACTAC TTTCAAATGG 1380
 CCAGAGCAAA CTGCAAGAAA TTGCTTTT TCTATATATC TTTTGCACAA TTTGAACCTG 1440
 CACAGGTGAA TGTCAAAAA AGTAAACAAC TTCTTCAAAA AGCTGTAGAA CTGGAGCAG 1500
 TACCACTAGA ATGCTGGAA ATGCCCCGCG GGAATTTAAA CCTCCAAAAA AAGCAGCTGC 1560
 TTTCAAGAGA GGAAGAAG AATTTATCAG CATCTACGGT ATTAACTGCC CAAGAATCAT 1620
 TTTCCGGTTC ACTGGGCAT TTACAGATA GGAACAACAG TTGTGATGCC AGAGGACAGA 1680
 CTACTAAGC CAGGTTTTTA TATGGAGAGA ACATGCCACC ACAAGATGCA GAAATAGGTT 1740
 ACCGGAATTC ATTGAGACAA ACTAACAAA CTAAACAGTC ATGCCCATTT GGAAGAGTCC 1800
 CAGTTAACC CTCAATGAG CCAGATTGTG ATGTGAAGAC AGATGATTCA GTTGACCTT 1860
 GTTTTATGAA AAGACAAACC TCTAGATCAG AATGCCGAGA TTTGGTTGTG CTTGGATCTA 1920
 AACCAAGTGG AATGATTCTC TGTGAATTAA GAAATTTAAA GTCTGTCAA AATAGTCAAT 1980
 TCAAGGACCC TCTGGTCTCA GATGAAAAGA GTTCTGAAT TATTATTACT GATTCAATTA 2040
 CCTGAAGAGA TAAACCGGAA TCAAGTCTTC TAGCTAAATT AGAAGAACT AAAGAGTATC 2100
 AAGAACCAAG GTTCCAGAG AGTAACAGGA AACAGTGGCA AGCTAAGAGA AAGTCAGAGT 2160
 GTATTAAACA GAATCCTGCT GCATCTTCAA ATCACTGGCA GATTCGGAG TTAGCCGAA 2220
 AAGTTAATAC AGAGCAGAAA CATACCACTT TTGAGCAACC TGTCTTTTCA GTTTCAAAAC 2280
 AGTCACCACC AATATCACA TCTAATGGT TTGACCCAAA ATCTATTGT AAGACCCAA 2340

GCAGCAATAC CTTGGATGAT TACATGAGCT GTTTTAGAAC TCCAGTTGTA AAGAATGACT 2400
 TTCCACCTGC TTGTCACTTG TCAACACCTT ATGGCCCAACC TGCCCTGTTTC CAGCAGCAAC 2460
 AGCATCAAACT ACTTGCCACT CCACTTCAAA ATTTACAGGT TTTAGCATCT TCCTCAGCAA 2520
 5 ATGAATGCAT TTCCGTTAAA GGAAGAATTT ATTCCATATT AAAGCAGATA GGAAGTGGAG 2580
 GTTCAAGCAA GGTATTTTCAG GTGTTAAATG AAAAGAAACA GATATATGCT ATAAATATG 2640
 TGAACCTAGA AGAAGCAGAT AACCAAACTC TTGATAGTGA COGGAACGAA ATAGCTTATT 2700
 TGAATAAACT ACAACAACAC AGTGATAAGA TCATCCGACT TTATGATTAT GAAATCACGG 2760
 ACCAGTACAT CTACATGGTA ATGGAGTGTG GAAATATGTA TCTTAATAGT TGGCTTAAAA 2820
 10 AGAAAAATC CATTGATCCA TGGGAACGCA ACAGTTACTG GAAAAATATG TTAGAGGCAG 2880
 TTCACACAAAT CCATCAACAT GGCATTGTTT ACAGTGATCT TAAACCAAGT AACTTTCTGA 2940
 TAGTTGATGG AATGCTAAGG CTAAATGATT TTGGGATTGC AAACCAAAATG CAACCAGATA 3000
 CAACAAGTGT TGTAAAGAT TCTCAGGTTG GCACAGTTAA TTATATGCCA CCAGAAACAA 3060
 TCAAGATAT GTCTTCTCTC AGAGAGAATG GGAATCTAA GTCAAAGATA AGCCCCAAAA 3120
 15 GTAGTGTG GTCTTAGGA TGTATTTTGT ACTATATGAC TTACGGGAAA ACACCATTTT 3180
 AGCAGATAT TATCAGATT TCTAAATTAC ATGCCATAAT TGATCCTAAT CATGAATTTG 3240
 AATTTCCCAA TATTCAGAGT AAAGATCTTC AAGATGTGTT AAAGTGTGTT TTAATAAGGG 3300
 ACCCAAAACA GAGGATATTC ATTCTGAGC TCCTGGCTCA TCCATATGTT CAAATTCAAA 3360
 CTCATCCAT TAACCAATG GCCAAGGAAA CCACTGAAAG AATGAATAT GTTCTGGGCC 3420
 20 AACTTGTG TCTGAATTTT CTTAACTCCA TTTTGAAAGC TGCTAAAACT TTATATGAAC 3480
 ACTATAGTGG TGGTGAAAGT CATAATTTCT CATCTTCCAA GACTTTTGAA AAAAAAGGG 3540
 GAAAAAATG ATTTGCAGTT ATTCGTAATG TCAGATAGGA GGTATAAAT ATATTGGACT 3600
 GTTATACTCT TGAATCCCTG TGGAAATCTA CATTTGAAGA CAACATCACT CTGAAGTGT 3660
 ATCAGCAAA AAAATTCAGT GAGATTATCT TTAAGAGAAA ACTGTAAAA TAGCAACCA 3720
 25 TTATGGCAGT GATATATTG TAGACTTGT TCTCTGTTT TATGCTCTTG TGTATCTAC 3780
 TTGACATCAT TTTACTCTTG GAATAGTGG TGGATAGCAA GTATATTCTA AAAAACTTG 3840
 TAAATAAGT TTTGAGGCTA AAATGA

30 A147 Protein sequence:
 Gene name: TTK protein kinase
 Unigene number: Hs.169840
 Probeset Accession #: M86699
 Protein Accession #: NP_003309
 Signal sequence: none found
 35 Transmembrane domains: none found
 Protein Kinase Domain: 510-775
 Cellular Localization: cytoplasmic and nuclear

40 1 11 21 31 41 51
 MNKVRDIRNK FNEDLTDEL SLNKISADTT DMSGTVNQIM MMANNPEDWL SLLKLEKNS 60
 VPLSDALINK LIGRYSQALE ALPPDKYQW ESFAIQVRF AELKAIQEPD DARDYFQMAR 120
 ANCKKPAFVH ISFAQFELSQ GNVKKSQQL QKAVERGAVP LEMLELALRN LNLQKQLLS 180
 45 REEKKNLAS TVLTQESFS GSLGHLQNRN NSCDNRGQT KARPLYGENM PFQDAEIGYR 240
 NSLRQNTKYK QSCPPGRVPV NLLNSPDCDV KYDSEVVPCE MKRQTSRSEC RDLVVPGEKP 300
 SGNDSCELRN LKSVQNSHFY EPLVSEKSS ELIITDSITL KNTKSSLLA KLEETKEYQE 360
 PVPFESNQQT WQAKRKSSCI NQNPAAASNH WQIPELARKV NTRQKHITFE QPVFVSQKS 420
 PPLSTBKWFD PKSICKTPSS NTLDDYMSCF RTPVVKNDFF PACQLSTFYG QPACFQQQOH 480
 50 QILATPLQL QVLASSANE CISVKRIYS ILKQIGSGGS SKVPQVINEK KQIYAIVYN 540
 LBEADNQTL DSYRNEIAYLN KLQHSQDKII RLYDYBITDQ YIYMVMCEGN IDLNSWLKK 600
 KSIDPWERKS YKKNMLEAVH TIHQHGVHVS DLKPFANFLIV DGMKLIDFG IANQMPPDT 660
 SVVDQSQVGT WYMPPEPAIK DMSSSRENGK SKSKISPKSD VWSGLCILYV MTKGTTFPQQ 720
 IINQISLEHA IIDPHEIEF PDIPKDLQD VLKCCLEKDP KQISIPELL AHPYVQIQTH 780
 55 FVMQMARGTT EEMKYVLGQL VQLNSPNSIL KAKTLYEHY SGGESENSSS SKTFEKKRKK 840
 K

60 A148 DNA SEQUENCE
 Gene name: endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
 Unigene number: Hs.258583
 Probeset Accession #: NM_012152
 Nucleic Acid Accession #: NM_012152
 Coding sequence: 43-1104 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51
 CTCTTTTAAA TTTCTTTCTA GGATGTTTAC TTCTTCTCCA CAATGAATGA GTGTCACTAT 60
 GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTGGA TGAAGTGGACA 120
 70 GGAACAAAGC TTGTGATGTG TTTGTGTGTT GGAAGTMTT TCTGCTGTT TATTTTTTTT 180
 TCTAATTTCT TGGTCATCGC GGCAGTGATC AAAAACAGAA AATTTCTATT CCCCTTCTAC 240
 TACCTGTGG CTAATTTAGC TGTCTGCGAT TTCTTGGCTG GAATTCGCTA TGTATTCTGT 300
 ATGTTTAA CAAGCCCACT TTCAAAAATC TTGACTGTCA ACCGCTGTT TCTCGCTCAG 360
 GGGCTTCTGG CAGTGTACTT GACTGCTTCC CTCACCAACT TGCTGGTTAT CGCCGTGGAG 420
 75 AGGCACATGT CAATCATGAG GATGCGGTC CATAGCAACC TGACCAAAAA GAGGGTGACA 480
 CTGCTCATTT TGTCTGTCTG GGCATCGGCC ATTTTATAGG GGGCGSTCCC CACACTGGGC 540
 TGAATTTGCC TCTGCAACAT CTCTGCTTGC TCTTCCCTGG CCCCATTATA CAGCAGGAGT 600
 TACCTGTTT TCTGGACAGT GTCCAACTC ATGGCCTTCC TCATCATGGT TGTGGGTGAC 660
 CTGCGGATCT ACGTGTACGT CAAGAGGAAA ACCAAGCTCT TGTCTCCGCA TACAAGTGGG 720
 80 TCCATCATGT GCTCGAGGAC ACCCATGAAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG 780
 GCGTTTGTGG TATGCTGGAC CCCGGGCTG GTGTTCTGCT TCTTCAAGCG CCTGAATGTC 840
 AGGCAGTGTG CGGTGACAGA TGTGAAAGG TGGTTCTGCT TGCTGGCGCT GCTCAACTCC 900
 GTGCGAACC CCACTATCTA CTCTACAAG GACGAGGACA TGTATGGCAC CATGAAGAAG 960
 ATGATCTGCT GCTTCTCTCA GGAGAACCCA GAGAGGCGTC CCTCTGCGAT CCCCTCCACA 1020
 GTCTCAGCA GAGGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080

GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCCTCTC GGCCCAACCA GGTGATGACT 1140
GTCTTAGG

5

A149 Protein sequence:

Gene name: endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
Unigene number: Hs.258583
Probeset Accession #: NM_012152
Protein Accession #: NP_036284
Signal sequence: none found
Transmembrane domains: 31-53, 66-88, 150-172, 190-211, 239-261, 277-295
Cellular Localization: plasma membrane

10

15

1	11	21	31	41	51	
MNECHYDKHM	DDFFYNRSNTD	TVDDWTGTLK	VIVLCVGTFF	CLFIFFSNEL	VIAAVIKNRK	60
FHFPPYYLLA	NLAADFFPAG	IAYVFLMFNT	GPVSKTLTVN	RWFLRQGLLD	SSLTASLTNL	120
LVIIVSRHMS	IMRMVHGNL	TKKRVILLIL	LWVAIAIFMG	AVPTLGWNC	CNISACSSLA	180
PIYSRSYLVF	WTVSHMAFL	IMVVVYLRIY	VYVKRKINVL	SPHTSGSISR	RRTPMKIMKT	240
VMTVLGAFVV	CWTPGLVVLL	LDGLNCRQCG	VQHVKNWFL	LALLNSVVP	IISYKDEDM	300
YGTMRKMICC	FSQENPERRP	SRIPSTVLRS	SDTGSQYIED	SISQGAVCNK	STS	

25

ProstateA150 DNA SEQUENCE

Gene name: ESTs
Unigene number: Hs.293616
Probeset Accession #: AW043782
Nucleic Acid Accession #: none found
Coding sequence: 38-1075 (underlined sequence corresponds to start and stop codon)

30

35

1	11	21	31	41	51	
AGCAACGACG	CGGGGACGCG	GGAGCGGCGG	CGGCGCCATG	TGGCTGCTGG	GGCCGCTGTG	60
CTGCTGCTGT	AGCAGCGCCG	CGGAGAGCCA	GCTGCTCCCG	GGGAACAAC	TCACCAATGA	120
GTGCAACATA	CCAGGCAACT	TCATGTGCAG	CAATGGACGG	TGCATCCCGG	CGCGCTGGCA	180
GTGTGACGGG	CTGCCGAGCT	GCTTCGACAA	GAGTGAIGAG	AAGGAGTGCC	CCAAGGCTAA	240
GTGGAATGTY	GGCCCAACCT	TCCTCCCTCG	TGCCAGCGGC	ATCCATTGCA	TCATGGGTGG	300
CTTCGGGTGC	AATGGGTGTC	AGGACTGTCC	CGATGGCAGC	GATGAAGAGA	ACTGCACAGC	360
AAACCCCTCT	CTTTGCTCGA	CCGCCCGCTA	CCACTGCAGG	AACGGCCCTC	GTATTGACAA	420
GAGCTTCATC	TGCGATGGAC	AGAATAACTG	TCAGACAAAC	AGTGATGAGG	AAAGCTGTGA	480
AAGTTCICAA	GAAACCGGCA	GTCGGCAGGT	GTTTGTGACT	TCAGAGAAC	AACCTGTGTA	540
TTACCCGACG	ATCACCTATG	CCATCATGGG	CAGCTCCGTC	ATTTTGTGTC	TGGTGGTGGC	600
CTGTCTGGCA	CTGGCTGTGC	ACCACGACGG	GAGGCGGAAC	AACCTCATGA	CGCTGCCCGT	660
GCACCGGCTG	CAGCACCCCG	TGCTGTCTGC	CCGCTCGGTG	GTCTTGGACC	ACCCCAACCA	720
CTGCAACGTC	ACCTACAAAG	TCAATATATG	CATCCAGTAT	GTGGCCAGCC	AGGCGGAGCA	780
GAATGCTGCG	GAGTAGGCT	CCCAACCTTC	CTACTCCGAG	GCTTGTCTGG	ACCAGAGGCC	840
TGGGTGGTAT	GACCTTCCTC	CACCGCCCTA	CTCTTCTGAC	ACGGAATCTC	TGAACCAAGC	900
CGACCTGCCC	CCCTACCGCT	CCCGGTCCGG	GAGTGCCAAC	AGTGCACAGT	CCGAGGCAGC	960
CAGCAGCTCG	CTGAGCGCTG	AAGACACCA	CCACAGCCCG	GGGCGCCCTG	GCCTCCAGGA	1020
GGGCACTGCT	GAGCCAGGCG	ACTCTGAGCC	CAGCCAGGCG	ACTGAAGAAG	TATAGTCCCG	1080
AGTTATTCCA	AGTCCATAT	GGGTTAACT	GCTCTGACTT	GTGGCCATTC	TACCAATTGG	1140
TGCTCAATGG	AAGCTCTTTA	AGCACCTGTA	AGGATGTCTC	AAGTTACAGT	TTGGGATATT	1200
AACATATCTT	GCATTCCTCT	CCTCCCGCAG	ACTTCAGAGA	TGTTTTCTTG	GGCTCTCAGT	1260
TGACATCTTC	TGTTGTGGCT	CTTTCTGTC	AGGTCACTCT	TCCCTTGGGA	CCCGAGATCA	1320
CACCTCATTT	TTTCACATTA	TTCTGTTTCT	GTGGGAGAGA	CAGCATATAA	AACAGTATTG	1380
AAATAGGCTG	GGAGAGAGCA	ATGTTTCTGT	GCTATATTGG	ATGCTCAGAA	GTGCAGGAGA	1440
CGCTGGACCC	AATTCTCTCT	GCTGGGTAGT	TACCTTAATG	CATTGGGGGA	TTTGGGTTAG	1500
ATGATCTAAC	CAGGAGGCCA	TCACTGGATG	GTCACCCCGC	CAAAAAAATT	CCATTGAGGC	1560
ATCAAAACCT	GCTTTGCACA	ATCCATTATT	ATGCCCCCAG	TTGAGCAGAG	TCAGTGGCCA	1620
AAGAAAACCT	TGAGACGTGAG	TAACACCTTT	CAGCAGTCCG	AACGTTATTT	TGGTTTGTGT	1680
AAGGACCTCG	AAACCATCTA	CCCTGTATTA	ATCTGGCTTT	TAGAAAATTG	CCCAAGAAATG	1740
CTCATTTCTA	GAGCTTTCTT	CAGCAGCATA	TATCATCAGC	CTCATCCTAA	AATAGGCAGG	1800
GAGCCCTTCC	CATGAGTTTA	TCCAAGTTCT	CAGCTCCTAA	AATGCAGGCT	GCCAGAGACC	1860
TACACCTGCC	CTGGCTCTAC	AGCCACTTAC	CTGGTTTCTG	GACTGTCAAC	CTCCAGCTGT	1920
ACCTGCCCGT	AGCCAGGGA	TGAGGAOCTA	ACTTGAGTTG	GGCCAAAGTC	TGACCTGGCT	1980
GTATGTCCCT	GTGGCCGACA	CCGAGCCGCT	CTGTCTATT	CATGCAGGCT	CAACACTGGC	2040
CTCCAAAGTT	CCCTTAACAC	TTGCAAGATC	CTTTTAACT	GTGCATTGAG	ACTTGAGGAC	2100
ACTGGTTTCT	ATCACAGGTC	AGAGCCATGT	TCAATACCTC	CAGCAAGCTC	TCTTGGCTCC	2160
CTGCACTGTG	CAGCTCTCTC	TTCCCAAGGT	CCCAATACCA	GCACCTCTAG	TTAGAGTTAG	2220
GGTCAGGCTC	AGGCTCTCTC	CAACATCCCA	GTAGTTTCTC	CTCTGAGACA	CATGGGCAAG	2280
AGACAAATTG	GAGTCAAGAT	TTTCCATTGT	GATCTATTTT	AAATCTTTTA	GAAATGCATT	2340
TGAAACAGTG	TGTTTGTGTT	TTCCCTTCTA	GTTAAGGGAC	TATTTATATG	TGTATAGGAA	2400
AGCTGTCTCT	TTTTTTGTTT	TTCTTTTAAC	AAGGTCCAAA	GAAAGATGCA	AAAGGAGATC	2460
ACACCTCTGC	CCGCTGTAGC	CCCGTGATAA	CAAGTCACTC	CAGACTAACC	TGTGTGCCAG	2520
ACATTGTGTC	ATTGTGTCAC	TTTGAGGTTA	TTATTATCA	AGTTCTTGAA	GGAGGAGGAA	2580
AGAGGGAGCT	CTCTCTCCCT	CCGTGTATAG	TCTCTATGTT	TGTGTAGTGT	TTTCTTTTCT	2640
TTCTCTGTGT	CCAGTCAAGC	ACAGGGCCCG	CCTCCCTGCA	GGAAATAGGG	GTAACACGTT	2700
AGGTGTGTGT	TGGCAAGAAA	CCACACTGAC	TGATGAGGGG	TAAATGGGAA	CCAGGTAGAG	2760
CCATCCGCGG	CAGCTGTAC	CCATTCAAG	CTTCTTTCCG	CAGCTGAAGA	AATGTTCAAT	2820

5 AACCTGTTTG ACCTAATTA AACACAGGCC TGCAGGAAGT GGGGCTAAAG TGGCATTGAG 2880
 TGATCCTGTG CTGTAGACTT TTCTTTCTTT TTTTAACCAA ATCCAAAGGA TGTACAGAA 2940
 AACCTAGCCA CTGGTATTTT GTTTTGTITA AAAAAAAGAA GAAAGAAAGA AAGAAAGAAA 3000
 AACGGAAAGG AACCTAGCTG CCTGTATCTT TCATTTTAA AATAGCACTT GAGTTATTTT 3060
 CTGAGTAATC CAATAAGGAA CTTTGTATGA CAGCCAGAAAT GTGTTAGAAC TCTGGCTGAA 3120
 CATTTTCATCT CCTGTGAGTC AGAAGGGGCTT TATTTCTCCC TTGTATGGGG CCCCTTCTTC 3180
 TTCTGTGTGC TCTGGAAGTT GTTTAGAGGA AAGAAATCTA ATTTTAATTA ATTGCGCAGT 3240
 GAGTTAATCT CACTCGCTTT TCTGCTTCCA GGCATCTTAG GAAACACAAA TGGTTTATGT 3300
 10 AGATAAGGGA TGCCCTACTAA TGCITTTTAA AAACAAACAG GACATTTTAT ATTATAGATT 3360
 TGATTTTATT AATGAATGTT TTTAAAAATA TATAAATAGG ACACCAAGAG GGCAGGGTTT 3420
 TTTTGGGGG GAGGGGGTTT GTTTTCCAAC TCAAGATGGC ACATTAGTGG CCAGCAATAT 3480
 TTTTAACTC ATTCCAACCA GGAAGCTTTT TTATACATTG CCTAAATCTA CGCCAAACAG 3540
 AAAATAGTCT CATCTCTTTT TTTCTCAAT GAGATCCGTG TTTTATTTTA GCATTAAATT 3600
 15 AGTTACACTG TGATGACTGG CCTATTACCT GACTCAGCTC CCTCTACCTT GAAATTGACA 3660
 TTTTAAAAA ATGCAACTAA GTGGTAAATA GTGTGTGACG CTCAAAGTTA ATGTAAACTG 3720
 GAAAGGTGTG GTGTGTGTGC TTTTGTGTGT TTGGTTAGGC TTGGTTTGTG TTTTAAATTT 3780
 TTATACCTTC TAATAAATTT GCAGTTTCAT TCTTTCTGTT TGTGCAAAAG GWMCTAMARM 3840
 AAMMAAAAC AHWTTTGGGG GGGCTTGGGC CTGGAAGAAA GTTTTAAACA CCCTTCGGG 3900
 20 TGGGGCGGG GGGCTTACCT AGGTACGGCG ACCACGGGGG CCAACCGGG ACCCCAGAG 3960
 GAAACCTCTG CCAAGAAAAA GGTGGCGAGA ATTCTCCACA CCAGAAAAA ACGGCGCGGG 4020
 GGAACCGCA GAGTGTTCG TAAACACAC CCGAAGAGAG AACTCAGAG CACACAAGCG 4080
 GGACTCAACC AGGAGGACCC AAGGGAACCC GATAGATTAC G

25 A151 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.293616
 Probeset Accession #: AW043782
 Protein Accession #: none found
 30 Signal sequence: 1-17
 Transmembrane domains: 169-191
 LDLA domains: 28-66, 70-108, 112-149
 Cellular Localization: plasma membrane

35
 1 11 21 31 41 51
 | | | | |
 MWLLGLPLCLL LSSAASQLL PGNFNECN IPGNFMCNNG RCIPGAWQCD GLPDCFDKSD 60
 40 EKCEPKAKSK CGTFFPCAS GIHCLIGRFR CNGFEDCPDG SDEENCTANP LLCSTARYHC 120
 KNGLCIDKSF ICDDQNNQCD NSEDESCESG QEPGSGQVVFV TSENQLVYYP SIXYAIIGSS 180
 VIFVLVALL ALVLHQRKR NMLTLFVHR LQHPVLLSRL VVLDHPHECN VTYNVNNGIQ 240
 YVASQASQNA SEVGSPPSY SALLDQRPAP YDLPPPPYSS DTESLNQADL PFYRSRSGSA 300
 NSASSQAASS LLSVEDTSHS PQQPGQESG AEPRDSEPSQ GTZEV

45 A152 DNA SEQUENCE

Gene name: Transmembrane protease, serine 2 (TMPRSS2)
 Unigene number: Hs.105807
 Probeset Accession #: T48536
 50 Nucleic Acid Accession #: NM_005656.1
 Coding sequence: 57-1535 (underlined sequences correspond to start and stop codons)

55 1 11 21 31 41 51
 | | | | |
 GTCATATTGA ACATTCCAGA TACCTATCAT TACTCGATGC TGTGTATAAC AGCAAGATGG 60
 CTTTGAATCT AGGGTCACCA CCAGCTATTG GACCTTACTA TGAAGACCAT GGATACCAAC 120
 CGGAAACACC CTATCCCGCA CAGCCCACTG TGGTCCCCAC TGTCTACGAG GTGCTACCGG 180
 CTAGTACTA CCGGTCCCCC GTGCCCCAGT ACGCCCCGAG GGTCTCTGAC CAGGCTTCCA 240
 60 ACCCGCTGCT CTGCAACGAG CCAAAATCCC CATCCGGGAC AGTGTGCACC TCAAAGACTA 300
 AGAAGACACT GTGCATCACC TTGACCTTGG GGAOCTTCCT CGTGGGAGCT GCGCTGGCCG 360
 CTGGCCTACT CTGGAAGTTC ATGGGCAGCA AGTGTCTCAA CTCTGGGATA GAGTGGGACT 420
 CCTCAGGTAC CTGCATCAAC CCCTCTAATC GGTGTGATGG CGTGTACAC TGCCTCCGGC 480
 GGGAGGACGA GAATCGGTGT GTTCGCCCTCT ACGGACCAAA CTTCATCCTT CAGATGTACT 540
 65 CATCTCAGAG GAAGTCTCTG CACCTCTGTG GCCAAGCGA CTGGAACGAG AACTACGGGC 600
 GGGCGGCTG CAGGGACATG GGCTATAAGA ATAATTTTAA CTCTAGCCAA GGAATAGTGG 660
 ATGACAGCGG ATCCACCAGC TTTATGAAC TGAACACAG TGCCGGCAAT GTCTGATCT 720
 ATAAAAAAT GTACCAAGT GATGCCTGTT CTTCAAAAGC AGTGGTTTCT TTACGCTGTT 780
 TAGCTGTGGG GGTCAACTTG AACTCAAGCC GCCAGAGCAG GATCGTGGGC GGTGAGAGCG 840
 70 CCGTCCCGGG GGCCTGGCCC TGGCAGGTCA GCCTGACGCT CCAGAGGCTC CAGTGTGCG 900
 GAGGCTCCAT CATCAACCCC GAGTGGATCG TGACAGCGGC CCACTGCGTG GAAAAACCTC 960
 TTAACAATCC ATGCAATTGG ACGGCATTTG CCGGGATTTT GAGACATCT TTCTGTTCT 1020
 ATGGAGCGGG ATACCAAGTA CAAAAGTGA TTTCTCTACC AAATTATGAC TTCAAGACCA 1080
 AGAACAAATG CATTCGCTG ATGAAGCTGC AGAAGCCTCT GACTTTCAAC GACCTAGTGA 1140
 75 AACCAATGTC TCGCCCAAC CCAGGCATGA TGTGACGCC AGAACAGCTC TGCTGATTT 1200
 CCGGTGGGG GGCACACGAG GAGAAAGGGA AGACCTCAGA AGTGTGAAC GCTGCCAAG 1260
 TGCTTCTCAT TGAGACACAG AGATGCAACA GCAGATATGT CTATGACAC CAGTACAC 1320
 CAGCCATGAT CTGTGCGGGC TTCTGTGAGG GGAAGCTGGA TTCTTCCAG GGTGACAGTG 1380
 GAGGGCCTCT GATCACTTCG AACACAATA TCTGGTGGCT GATAGGGGAT ACAAGCTGG 1440
 80 GTTCTGGCTG TGCCAAAGCT TACAGACCAAG GAGTGTACGG GAATGTGATG GTATTCAAG 1500
 ACTGGAATTA TCGCAATATG AAGGCAAAAG GCTAATCCAC ATGCTCTTGG TCTTGAAGT 1560
 CGTTTACAA GAAACCAATG GGGCTGGTTT TGCTTCCCGG TGCATGATTT ACTCTTAGAG 1620
 ATGATTCAGA GGTCACTTCA TTTTATTTAA ACAGTGAAGT TGTCTGGCTT TGGCACTCTC 1680
 TGCCATACAG TGCAGGCTGC AGTGGCTCCC CTGCCAGGCC TGCTCTCCCT AACCCCTGT 1740

5 CCGCAAGGGG TGATGGCCGG CTGGTTGTGG GCACTGGCGG TCAATTGTGG AAGGAAGAGG 1800
 GTTGGAGGCT GCCCCCATTT AGATCTTCCT GCTGAGTCCT TTCCAGGGGC CAATTTTGGA 1860
 TGAGCATGGA GCTGTCACTT CTCAGCTGCT GGATGACTTG AGATGAAAAA GGAGAGACAT 1920
 GGAAGGGGAG ACAGCCAGGT GGCACCTGCA GCGGCTGCCC TCTGGGGCCA CTGGGTAGTG 1980
 TCCCCAGCCT ACITTCACAG GGGATTTTGC TGATGGGTTC TTAGAGCCCT AGCAGCCCTG 2040
 GATGGTGGCC AGAAATAAAG GGACCAAGCC TTCATGGGTG GTGACGTGGT AGTCACTTGT 2100
 AAGGGGAACA GAAACATTTT TGTCTTATG GGGTGAGAAT ATAGACAGTG CCTTGGTGTC 2160
 GAGGGAAGCA ATTGAAAGG AACTTGCCTT GAGCACTCCT GGTGACAGGT TCCACCTGCA 2220
 CATTGGGTGG GGCCTCTGGG AGGAGAGACT AGCCCTCCTC CTCATCCTCC CTGACCCCTG 2280
 10 TCCTAGCACC CTGGAGAGTG AATGCCCTT GGTCCCTGGC AGGCGGCCAA GTTTGGGCACC 2340
 ATGTCGGCCT CTTCAGGCCT GATAGTCATT GGAATTTGAG GTCCATGGGG GAAATCAAGG 2400
 ATGCTCAGTT TAAGGTACAC TGTTTCCATG TTATGTTTCT ACACATTGAT GGTGGTGACC 2460
 CTGAGTTCAA AGCCATCTT

A153 Protein sequence:

Gene name: Transmembrane protease, serine 2 (TMPRSS2)
 Unigene number: Hs.105807
 20 Probeset Accession #: T48536
 Protein Accession #: NP_005647.1
 Signal sequence: none found
 Transmembrane domains: 85-107
 Cellular Localization: plasma membrane

1 11 21 31 41 51
 | | | | |
 MAINSGSPFA IGFYIENHGY QPENFYPAQP TVVPTVYEVH PAQYYPSPVP QYAPRVLTA 60
 30 SNPVVCTQPK SPSTGVTCTSK TKKALCITLT LGTFLVGAAL AAGLLNKEFG SKCSNGSIEC 120
 DSSGTCINPS NWCQGVSHCP GGDENRCVR LYGPNFILQM YSSQRKSWHP VCQDDWNNENY 180
 GRAACRDMGY KNNFYSSQGI VDDSGSTSPM KLNTSAGNVD IYKLYHSDA CSSKAVVSLR 240
 CLACGVNLNS SRQSRIVGGE SALFGAMPWQ VSLRVQNVHV CGGSIITPEW IVTAACRVEK 300
 PLNTPWHNTA FAGILRQSEFM FYGAGYQVQK VISHPNYDSK TKNNIDIALMK LQKPLTFNDL 360
 VKFVCLFNPG MNLQPEQLCW ISGWGATEEK GKTSEVLNAA XVLLEIETQRC NSRYVVDNLI 420
 35 TPAMICAGFL QGNVDSQCGD SGGPLVTSNN NIWWLIGDTS WSGSCAKAYR PGVYGNVMVF 480
 TDWYYQMKR NG

A154 DNA SEQUENCES

Gene name: Homo sapiens cDNA FLJ13581 fis, clone PLACE1009039
 40 Unigene number: Hs.129179
 Probeset Accession #: A1694767
 Nucleic Acid Accession #: A1694767
 Coding sequence: 130-1086 (underlined sequences correspond to start and stop codons)

45 1 11 21 31 41 51
 | | | | |
 CAGAGAGGCT GTATTTCACT GCAGCCTGCC AGACCTCTTC TGGAGGAAGA CTGGACAAAG 60
 GGGGTCAACAC ATTCTCTCCA TACGGTTGAG CCTCTACCTG CCTGGTGCIG GTCAACAGTTC 120
 50 AGCTTCTTCA TGATGGTGGG TCCCAATGGC AATGAATCCA GTGCTACATA CTTTCTCTTA 180
 ATAGGCCCTCC CTGGTTTAGA AGAGGCTCAG TTCTGGTTGG CCTTCCCATT GTGCTCCCTC 240
 TACCTTATTT CTGTGCTAGG TAACCTGACA ATCATCTACA TTGTGCGGAC TGAGCAGAGC 300
 CTGCATGAGC CCATGTATAT ATTCTTTTGC ATGCTTTTCA GCATTGACAT CCTCATCTCC 360
 ACCTCATCCA TGCCCAAAAT GCTGGCCATC TTCTGGTTCA ATTCCACTAC CATCCAGTTT 420
 55 GATGCTTGTC TGCTACAGAT GTTGGCCATC CACTCCTTAT CTGGCATGGA ATCCACAGTG 480
 CTGCTGGCCA TGGCTTTTGA CCGCTATGTG GCCATCTGTC ACCCACTGCG CCATGCCACA 540
 GTACTTACGT TGCCTCGGGT CACCAAAATT GGTGTGGCTG CTGTGGTGGG GGGGGCTGCA 600
 CTGATGGCAC CCTTCTCTGT CTTCTACAGC CAGCTGCCCT TCTGCGGCTC CAATATCTTT 660
 TCCCATTCCT ACTGCTTACA CCAAGATGTC ATGAAGCTGG CCTGTGATGA TATCCGGGTC 720
 60 AATGTGCTAT TGGCTCTTAT GGTCTATC TCCTCCATTC GCTTGGACTC ACTTCTCATC 780
 TCCTTCTCAT ATCTGCTTAT TCTTAAGACT GTGTGGGCT TGACACGAGA AGCCCAAGGC 840
 AAGGCATTG GCATTTCCGT CTCTCATGTG TGTGCTGTGT TCATATTCTA TGTACCTTTC 900
 ATTGATTTGT CCATGGTGCA TCCTTTTATC AAGCGGGTGT ACTCTCCACT GCCGTGATC 960
 TTGGCCAATA TCTATCTGCT GGTTCCTCTT GTGCTCAACC CATTTGTCTA TGGAGTGAAG 1020
 65 ACAAAGGAGA TTCCACAGCG CATCCTTGA CTTTTCCATG TGGCCACACA CGCTTCAGAG 1080
 CCTTAGGTGT CAGTGATCAA ACTTCTTTTC CATTCAGAGT CCTCTGATTC AGATTTTAAT 1140
 GTTAACATTT TGAAGACAG TATTCAGAAA AAAAATTTCC TTAATAAAAA TACAACCTAG 1200
 ATCCTTCAAA TATGAACATG GTTGGGGAAT CTCCATTTT TCAATATTAT TTTCTTCTTT 1260
 GTTTTCTTGC TACATATAAT TATTAATACC CTGACTAGGT TGTGGTTGGA GGGTTATTAC 1320
 70 TTTTCATTTT ACCATGCAGT CCAATCTTAA ACTGCTTCTA CTGATGGTTT ACAGCAITTC 1380
 GAGATAAGAA TGGTACATCT AGAGAACATT TGCCAAAGGC CTAAGCACAG CAAAGGAAAA 1440
 TAAACACAGA ATATATATAA ATGAGATAAT CTAGCTTAAA ACTATAACTT CCTCTTCAGA 1500
 ACTCCCAACC ACATTTGATC TCAGAAAAAT ACTGTCTTCA AATAGACTTC TACAGAGAAC 1560
 75 AAATAATTTT TCCTCTGGAC ACTAGCACAT AAGGGGAAGA TTGGAAGTAA AGCCTTGAAA 1620
 AGAGTACATT TACCTACAGT AATGAAAGTT GACACACTGT TCTGAGAGTT TTCAAGCAT 1680
 ATGGAACCTG TTTTCTCTAT TTAATTTTCT TATCAACCTT TTAATTAGGC AAAGATATTA 1740
 TTAGTACCTT CATTTAGACC ATGGGAAAAA TGATGTTTCA TGGGGATCAG TGAATTAAT 1800
 GGGGTCTATC AAGTATAAAA ATTAATAAAA AAAGACTTCA TGCCCAATCT CATATGATGT 1860
 80 GGAAGAATCT GTTAAAGAGC CAACAGGGTA GTGGGTTAGA GATTTCCAGA GTCTTACATT 1920
 TTCTARAGGA GGTATTTAAT TTCTTCTCAC TCATCCAGTG TTGATTTTAT GAATTTCTGT 1980
 GCAACAGAAC TCATGGCTTT AATCCCACTA GCTATTGCTT ATGTGCTGG TCCAATTGCC 2040
 AATTACCTGT GTCTTGGAG AAGTGATTTT TAGGTTTACC ATTATGGAAG ATCTTTATTC 2100
 AGAAGTCTG CATAGGCTTT ATTTATTTT AAAAGTTCCA TAGGGTGTTC 2160
 TGATAGGCAG TGAGGTTAGG GAGCCACCAG TTATGATGGG AAGTATGGA TGGCAGGTGT 2220
 TGAAGATAAC ATTGGCCTTT GAGTGTGAC TCGTAGCTGG AAGTGAAGG AATCTTCAG 2280

5
10
15

```

ACCATGCTTT ATTGGGGCT TTGTGCAGTA TGGAAACAGG ACITTTGAGAC CGGGAAGCA 2340
ATCTGACTTA GGCATGGGAA TCAGGCATTT TTGCTTCTGA GGGGCTATTA CCAAGGGTTA 2400
ATAGGTTTCA TCTTCAACAG GATATGACAA CAGTCTTAAC CAAGAACTC AAATTACATA 2460
TACTAAACA TGTGATCATA TATGTGGTAA GTTTCATTTT CTTTTCATC CCTCAGGTTT 2520
CCTGATATGG ATTTCCTATNA CATGCTTTCA TCCCCTTTTG TAATGGATAT CATATTGGGA 2580
AATGCCTATT TAACTACTGT ATTGTGCTGT GGACTGTAGG CCCATGAGGG CACTGTTTAT 2640
TATTGAATGT CATCTCTGTT CATCATTGAC TGCTCTTTGC TCATCATTGA ATCCCCCAGC 2700
AAAGTGCCTA GAACATAATA GTGCTTATGC TTGACACCGG TTATTTTTC TCAAACTCTA 2760
TTCCTTCTGT GCTGAACACA TAGCCAGGCA ATTTTCCAGC CTTCTTTGAG TTGGGTATTA 2820
TTAAATTTTA GCCATTACTT CCAATGTAGG TGGAAAGTGC ATGTGCAATT TTTATACCTG 2880
GCTCATAAAA CCCTCCCATG TGCAGCCTTT CATGTTGACA TTAATGTGA CTGGGAAGC 2940
TATGTGTAC ACAGAGTTAA TTAACCGAA AGGCTGGNA ATTTTGGNN AANNAACTG 3000
TGGCCNNGAG GCCNCNCAAC CTTTNNNA ATTTGGCAAN NTCOCACCTT GTANTTTGGT 3060
AAGGAGGCCA GTTGATAAG TGAATAATA AGTACTATTG TGTC

```

A155 PROTEIN SEQUENCE

20
25

```

Gene name: Homo sapiens cDNA FLJ13581 fis, clone PLACE1009039
Unigene number: Hs.129179
Protein Accession #: not available
Signal sequence: none found
Pfam domain: 7tm_1 [43-293]
Transmembrane domains: 29-51, 57-79, 82-104, 203-225, 239-261, 273-295
Cellular Localization: not determined

```

30
35

```

1 11 21 31 41 51
| | | | |
MVDPNGNESS ATYFILIGLP GLEEAQFWLA FPLCSLYLIA VLGNLTIIYI VRTEESLHEP 60
MYIFLCMLSG IDILISTASM PKMLAIFWFW STTIQFDACL LQMFHHSLS GMESTVLLAM 120
AFDRYVAICH PLRHATVTL PRVTKIGVAA VVRGAALMAP LRVFIKQLFF CRENILSESY 180
CLEQDVMLKA CDDIRVNVVY GLIVIIISAIG LDSLLISPSY LLILKTVLGL TREQAQAFG 240
TCVSHVCAVF IFYVFFIGLS MVHRFSKRRD SPLFVILANI YLLVFPVLNP IVYGVTKFI 300
RQRILRLPHV ATHASEP

```

A156 DNA SEQUENCE

40

```

Gene name: vasoactive intestinal peptide receptor 1
Unigene number: Hs.198726
Probeset Accession #: X77777
Nucleic Acid Accession #: NM_004624.1
Coding sequence: 57-1544 (underlined sequences correspond to start and stop codons)

```

45
50
55
60
65
70
75
80

```

TCGGAGCCTG CGGAGGGTGG TGGTGGTGGT GGTGGTGGCC CTGCGCCGCC TCACTCATGC 60
CTCTGCTGCC TCTGCTCTCG CTCAGGCGCC TCGGTGGCGG TTGGTGGCGG GTTACGCGGC 120
TGGTGGTGGC GCGGCGCGGG GCTGCTCTTC GGGGAGGCCG GGGCGGATCT CGCGCGCGAG 180
GCGGCGCGCG GCGAGGCTGG GTGCGCGCGG GAGGCGCGCT CGAGCTTCGT GCTGCGCGCT 240
CGCTCTTGGG CTCTGCTGCG CAGGAGGAGT GTGACTATGT CGAGATGATC GAGGTGCGAG 300
ACAAGCAGTG CCTGAGGAGG GCCCAGCTGG AGAATGAGC AATAGGCTGC AGCAAGATGT 360
GGGACACCTT CACCTGCTGG CCAGCCACCC CTGCGGCCCA GGTAGTTGTC TTGGCTGTGC 420
CCCTCATCTT CAAGCTCTTC TCCTCCATTC AAGGCCBCRA TGTAGCCGCG AGCTGCACCG 480
AGCAAGGCTG GAGGCACCTG GAGCCTGGCC CGTACCCCAT TGCCGTGGGT TTGGATGACA 540
AGGCAAGCAG TTTGGATGAG CAGCAGACCA TGTCTACCGG TTCTGTGAAG ACCGGCTACA 600
CCATTGGCTA CGGCTGTGCC CTGCGCACCC TTCTGGTGGC CACAGCTATC CTGAGCCTGT 660
TCAGGAAGCT CCATGTCACG CGGAACCTACA TCACCATGCA CCTCTCATCA TCCTCATGCC 720
TGAGGGCTGC GGCTGTCTTC ATCAAGACTT TGGCCCTCTT CGACAGCGGG GAGTGGGACC 780
AGTGTCCCGA GAGCTCGGTG GGCTGTAAAG CAGCCATGGT CTTTTCCTCA TATTGTGTCA 840
TGGCTAACTT CTTCTGGCTG CTGGTGGAGG GCCCTACCTT GTACACCCCTG CTTGCCGTCT 900
CCTCTCTCTC TGAGCGGAAG TACTTCTGGG GGTACATACT CATCGGCTGG GGGGTACCCA 960
GCACATTCAC CATGCTGTGG ACCATGCCCA GATCCATTT TGAGGATTAT GGTCTGTCTA 1020
GGTGTGGGGA CACCATCAAC TCCTCACTGT GGTGATCAT AAAGGGCCCC ATCTCTACCT 1080
CCATCTTGTT AAATTCATC CTGTTTATTT GCATCATCGG AATCCTGCTT CAGAACTGTC 1140
GGCCCCCAGA TATCAGGAAG AGTGACAGCA GTCCATFATC AAGGCTAGCC AGGTCCACAC 1200
TCCTGCTGAT CCCCCTGTTT GGAGTACACT ACATCATGTT CGCCTCTTIT CCGGACAAAT 1260
TTAAGCCTGA AGTGAAGATG GTCTTTGAGC TCGTCTGGGG GTCTTTCAGG GGTTTTGTGG 1320
TGGCTATCTT CTACTGCTTC CTCATAGGTG AGGTGCGAGC GGAGCTGAGG GGGAAAGTGG 1380
GGGCTGGGCA CCGTGGAGGC GTCTGGGGCT GGAACCCCAA ATACCGGCAC CCCTCGGBAG 1440
GCAGCAACCG CCGCACGTGC AGCAGGCAAG TTCTCATGCT GACCCGCGTC AGGCCAGGTG 1500
CCGCGCGCTC CTCCAGCTTC CAAGCCGAAG TCTCCCTGGT CTGACCAOCA GGATCCGAGC 1560
CCAGCGGGCC CTCCCGCCCC CTTCCTACTC GCACGACAGC CCGGGGACAG AGGCTGCCCC 1620
GGGCGCGCCA GCGCGGCCCC TGGGCTGGGA GGTGCGCCCC GGGCCCGCTG TCTCTGGTTC 1680
GGACACTCTT AGAGAACGCA GCCCTAGAGC CTGCTGGGAG CGTTTCTAGC AAGTGAGAGA 1740
GATGGAGACT CTTCTCTGCG AGGATGCGAG TGGAACTCAG TCATTAGACT CCTCTCCCAA 1800
AGGCCCGCTA CCGCAATCAA GGGCAAAAG TCTACATACT TTCTCTCTGA CTCGCGCCCC 1860
TGCTGCTCTT CTGCGCAAT TGGAGGAAG CAACCGGTGG ATCCTCAAAC AACACTGGTG 1920
TGACCTGAGG GCAGAAAGGT TCTGCGCGGG AAGGTACCCA GCACCAACAC CACGTTAGTG 1980
CTGAATATT CACCATTTGT GTCAAGTTC TTTGGGTAA GCATTACCA TCAGGCATTT 2040
GACTGAAGAT GCAGCTCACT ACCCTATTCT CTCTTTACGC TTAGTTATCA GCTTTTAAAA 2100
GTGGGTATT CTGAGGTTT TGTTTGGAGA GCACACCTAT CTTAGTGGTT CCCCACGAA 2160
GTGGACTGGC CCTGGGTCA GTCTGGTGGG AGGACGGTGC AACCCAGGA CTGAGGGACT 2220
CTGAGGCTCT TGGGAATGA GAAGGCAGCC ACCAGCGAAT GCTAGGTCTC GGCATAGGCC 2280
TACCTGCTCT CCAAGTCTCA GTGGCTTCAT CTGTCAAGTG GGACTCTGTC ACACGAGCCA 2340
TTCTATCTC TCTGTCTGT GGAAGCAACA GAATCAAGA GACTGCCCCC CTGTCCACC 2400
CACTATGTG CCAACTGTTG TAACTAGGCT CAGAGATGTG CACCCATGGG CTCTACAGA 2460

```

5
AAGCAGATCC TCACCCTGCT ACACATACAG GATTGAACT CAGATCTGTC TGATAGGAAT 2520
GTGAAAGCAC GGAATCTTAC TGCTAACTTT TGTGTATCGT AACGAGCCAG ATCCTCTTGG 2580
TTATTGTGTT ACCACTTGTA TTATTAATGC CATTATCCCT GAATTCCTCT TGCACCCCA 2640
CCCTCCCTGG AGTGTGGCTG AGGAGGCCTC CATCTCATGT ATCATCTGGA TAGGAGCCTG 2700
CTGGTCACAG CCTCTCTGT CTGCCCTTCA CCCGAGTGGC CACTCAGCTT CCTACCCACA 2760
CCTCTGCCAG AAGATCCCTC CAGGACTGCA ACAGGCTTGT GCAACAATAA ATGTTGGCTT 2820
GGAAAAAAA AAAA

10 A157 Protein sequence:
Gene name: vasoactive intestinal peptide receptor 1
Unigene number: Hs.198726
Probeset Accession #: X77777
Protein Accession #: JC2195
Signal sequence: none found
15 Transmembrane domains: 181-202, 214-236, 255-277, 290-311, 332-354, 377-399, 408-430
Cellular Localization: plasma membrane

20
1 11 21 31 41 51
MPPPPLLSLR RLGGGWSAVT RLVVAAAGAR SRGGRGSGRG AGGGRGGVA RRRRLRLRAA 60
RSLLGSSLOE RCDYVQMIEV QHKQCLBEAQ LENETIGCSK MWDNLTCWPA TPRGQVVVLA 120
CPILIFKLFS IQGRNVSRSC TDEGNTHELEP GPYPIACGLD DKAASLDEQQ TMFYGSVKTG 180
YTTIGYGLSLA TLLVATAILL LFRKLHCTRN YIHMHLFISF ILRAAAVFIR DLALFDSGES 240
25 DQCSGSGVGC KAAMVFPOYC VMAFFFWLLV EQLYLYTLA VSFFSERKYF WGYILIGWV 300
PSFTFMVWT ARIHFEDYGL LRCWDITNSS LHWIIRKPII TSILVNFILF ICIIRILLQK 360
LRPPDIRKSD SSPYRLARS TLLILFLPGV HYLMFAFFD NFKPEVKNVF ELVVGSEFQF 420
VVALLYCFLN GEVQAELERR WRRNHLQGVL GWNPKYRHPS GGSNGATCST QVSMLETRVSP 480
30 GARRSSSFQA EYSLV

A158 DNA SEQUENCE
Gene name: ESTs
Unigene number: Hs.29383
Probeset Accession #: AW207206
35 Nucleic Acid Accession #: AL133619
Coding sequence: 1-2070 (underlined sequences correspond to start and stop codons)

40
1 11 21 31 41 51
ATGAGCGGTG CGGGGGTGGC GCTGGGACG CGGCCCCCA GCTGCGGAC CCGGGGCTCT 60
CGGCGCCGGC GCCAGGCCCC CTCGTGGGCG GTCCAGTCTT TGAGGCGGCA GAGCCCGCAG 120
CTCAGGCAGA GCGACCCGCA GAAACGGAACT CTGACCTTGG AGAAAGCCCT GCAGTTCCTG 180
CAGCAGCAGC ACTCGGAGAT GCTGGCCAAAG CTCCATGAGG AGATCGAGCA TCTGAAGCGG 240
GAAACAGAG GTGAGCGGCG CGGGGGCCCT AGGCGGCCCC TGCCCTCCCA GGCACACTCA 300
45 ACACCTGCCG TCCCGCAGCA CAGAAACACA GCCATCACT CCAGCACACG CCTGGGCTCA 360
GGGGGAACAC AGGACGGGGA GCGCCCTCCAG ACTGTCTCTG CCCACCTGGC TGCACCTGGC 420
CCTGTATGCC AACCCAGTGG GTACAGGTTC TGGGGGACCT GGACAGATGC CGCTACCTCT 480
AGCCGTGGCT GGCAGATGTT ATGCAGCCAA GCACAGCACG TGCTGTCTCT GGGGAAGCCA 540
GGGCTCAGG TCATTGAGG GGGGAGGTG GCCACAGGCT GCTCCCTCAG CCTCCCTCTT 600
50 CCAAGTAGAG CTGAATGGG AAGGAACCTT TGGGACAGCC CCGCCCTGCG TAGATCTTTG 660
CCTCAGATTG CTGCTGTGGC CAGGCCCCAG ATTTCAGGCC CTATGGCTCT GAGTCTCTAC 720
ATGCTGGGGG CCGAGGGGAT ATGGACACAC TCCATCCAGG GATCCCTTCC TGCCATCTGG 780
GCAGCAACCA TGGGGAAGC AGAGTCTCTG TTCTTTCGCA CTGTCTCAAG 840
GCATCTCCCC ATCTCTACAG CCGCCCCCAG CCAGCCGAGG ATCCTTGGCT GTGTCTCAA 900
55 GCTCACTTCC CATTAFTCTT GGGGCTGGGG CTGACATCAG GAGGACATCT GACTGGTGGG 960
TGGAGCCAGC CTGGGAACAT CGCAGCTGGG GCAGTGCCTA GGGCTCTCC TTCCAGGGA 1020
GACATGTAGA AGGGGGTTGA GGGAGGGCCC TTCCCTAGCC GCTGTGGCAA CTCCAGTGAG 1080
CTCTCTGGG CAAAGTGTGG CCAAGTCTGG CAGCCCGAGC CCGTCACTGC TGGGGAGCCT 1140
GACAGACAC GGGAGAGGCG CATGCTTTC CTGGGACCT GCTGTCCAT GTGTCCCAAG 1200
60 CCTCCTGCT TTCCAGATGG CCGCTCAGGA AACCACTTT CAGGGGCTC TGCTCCCTTG 1260
GGCGCTCGCT GGTCTGCAAT CAACGAGTG TGGGTAGAGC CGGGAGGACC CAGCCCTGCC 1320
AGGCTGAAGG AGGGCTCCTC AGGACACAC AGGCCAGGAG GCAAGCGTGG GCGTCTTGG 1380
GGCGGTAGCG CCGACACTGT GCGCTCTCCT GCAGACAGCC TCTCCATGTC AAGCTTCCAG 1440
65 TCTGTCAAGT CCATCTCTAA TTCAGCCAAC TCTCAAGGCA AGGCCAGGCT CCAGCCCGGC 1500
TCTTCAACA AGCAAGATTC AAAAGCTGAC GTCTCCAGA AGGCCGACCT GGAAGAGGAG 1560
CCCTACTTTC ACAACAGCAA GCTGGACAAA GTTCTCTGGG TACAAGGGCA GGCAGAAAAG 1620
GAGGAAGCAG AGGCTCTTAA TGCAGGAGCT GCTGTATGAG GGAACAGCCA GCACAGGGC 1680
AGGCAGATGG GGGCGGGGGC ACACCCOCCA ATGATCCTGC CCCTTCCCT GCGAAAGCCC 1740
70 ACCACACTTA GCGAGTGGCA AGTGTCTATC CCGAGCTGT GGAATACCAA CCTCTGAGC 1800
ACCCAGAGCG TCGGGCACTT CAGTCCCTC CTGGAAGGGA GCCAGAGGCT CCAGGCAGCC 1860
CGGGAGGAAG CTAGCTTTCC CAGGGACCAA GAAGCCAGCG ATTTCCOCCA GGTCTCCACC 1920
AAGAGCTCTT CCAAGAAATG CCTGAGCCCA CCTGTGGCGG AGCGTGCCAT CCGCCCGCA 1980
75 CTGAAGCAGA CCGGAAGAA CAACCTTGCC GAGAGGCAGA AGAGGCTGCA GCAATGCAG 2040
AAACGCGGCC TGATGCTC AGTGTCTTGA

A159 Protein sequence:
Gene name: ESTs
Unigene number: Hs.29383
Probeset Accession #: AW207206
Protein Accession #: T43457
Signal sequence: none found
80 Transmembrane domains: 303-322
Cellular Localization: not determined

```

1      11      21      31      41      51
5      |      |      |      |      |
MSGAGVAAGT RPPSSPTPGS RRRRORPSVG VQSLRPQSPQ LRQSDPQKRN LDLEKSLQFL 60
QQQSEEMLAK LHBEIEHLKR ENKGEFARGP RPALPPQAKS TLPLPOHRNT AINSSTRIGS 120
GGTQDGEPLQ TVLAHLAALA PVCQPSGYRF WGTNTDAATS SRGWTMLCSQ AQHVLLSGSP 180
GPEVIAGRQV ATGCSPLPFP PSRAEMGRNP WDSPPCARSL PQIAAVARPR ISSPMALSPH 240
MLGAQGIWTH SIQGSLEPAIW AATMGTKGGS RVLPFFCHLSK ALPHPDSPGPH PAQDPGLWSQ 300
10      AHFPLSLGLG LTSGGHLTGG WSQPGNIAAG AVPRALPSQG DMEKGVGGFP FFSRCGNSEE 360
LFWAKCGPSR QPQPCSGADA DRTREEMLS LGTCCSMCPK PSCFPDGPSP NHLSRASAPL 420
GARWVCINGV WVEPPGGPSPA RLKEGSSRTH RPPGKRGRLA GGSADTVRSP ADSLSMSSFQ 480
SVKSISSNAN SQGKARPPQG SFNKQDSKAD VSQKADLEEE FLHNSKLDK VPGVQQQARK 540
EKAEASNAGA ACMGNSQHQG RQMGAGAHFP MILPLPLRKP TTLRQCEVLI RELWNTNLLQ 600
15      TQELHHLKSL LEGSQRPQAA PEEASFPRDQ EATHFPKVST KSLSKKCLSP PVAERAILPA 660
LKQTPKNFPA ERQKRLQAMQ KRRLHRSVL

```

A160 DNA SEQUENCE

Gene name: LIV-1 protein, estrogen regulated

Unigene number: Hs.79136

Probeset Accession #: U41060

Nucleic Acid Accession #: NM_012319.2

Coding sequence: 138-2405 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
30      |      |      |      |      |
CTCGTGC CGA ATTCCGGCAG AGACCCGCGTG TCGCGCCCTG GTAGAGATTT CTCGAAGACA 60
CCAGTGGGCC CGTGTGGAAC CAAACCTGCG CGCGTGGCCG GGCCTGGGGA CAACGAGGCC 120
GCGGAGACGA AGGCGCAATG CGGAGGAAGT TATCTGTAAT CTGTATCTGT ACCTTTGCCG 180
TCTCTGTAC AAATCCCTCT CATGAACTAA AAGCAGCTGC TTCCCCCAG ACCACTGAGA 240
AAATAGTACC GAATTGGGAA TCTGGCATTG ATGTGACCTT GGCATTTTCC ACACGSCAAT 300
ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAAATAA TTCTTTGTCA GTTGAGGGT 360
35      TCAGAAATTT ACTTCAAAT ATAGGCATAG ATAAGATTAA AAGAATCCAT ATACACCATG 420
ACCAACACCA TCACCTCAGC CACGAGCATC ACTCAGACCA TGAGCGTAC TCAGACCATG 480
AGCATCCTAC AGACACGAGC CATCACTCTG ACCATGATCA TCACCTCTAC CATATCATG 540
CAGCTTCTCG TAAAAATAAG CAAAAAGCTC TTTGCCCGAG CCATGACTCA GATAGTTCAG 600
GTAAAGATCC TAGAAACAGC CAGGGGAAAG GAGCTCAACC ACCAGAACAT GCCAGTGGTA 660
40      GAGGAGATGT CAAGACAGT GTTAGTCTA GTGAAGTGA CTCAACTGTG TACAACACTG 720
TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAATCTTTCC 780
CCAAAGATGT AAGCAGCTCC ACTCCACCCA GTGTCACTC AAGAGAGCCG GTGAGCCGCG 840
TGGCTGGTAG GAAACAAAT GAATCTGTGA GTGAGCCCCG AAAGGGCTTT ATGTATTCCA 900
GAAACACAAA TGAAATCTCT CAGGAGTGTG TCAATGCATC AAGCTACTG ACATCTCATG 960
45      GCATGGGCAT CCAGGTTCGG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA 1020
TCACCAAAAT TGATCTAGA TCTTGCTGTA TTCATACAG TGAAGAAGAG GCTGAAATCC 1080
TCCAAAGAC CTATTCAATA CAATAGCCCT GGGTTGGTGG TTTTATAGCC ATTTCCATCA 1140
TCAGTTTCTT GTCTCTGCTG GGGTTATCTT TAGTGCCCTC CATGAATCGG GTGTTTTTCA 1200
AATTTCTCTC GAGTTTCTCT GTGGCACTGG CGTTTGGGAC TTTGAGTGGT GATGCTTTTT 1260
50      TACACCTTCT TCCACATCTC CATGCAATGC ACCACCATAG TCATAGCCAT GAAGAACCAG 1320
CAATGGAATG GAAAGAGGA CCACCTTTTC GTCATCTGTC TTCTCAAAAC ATAGAAGAAA 1380
GTGCTTATTT TGATTCCACG TGGAGGGGTC TAACAGCTCT AGGAGGCCCTG TATTTCTATG 1440
TTCTTGTGTA ACATGTCTC ACATTGATCA AACAATTTAA AGATAAGAG AAAAAGAATC 1500
AGAGAAACCC TGAAATCTAT GATGATGTGG AGATTAAAG GCASTTGTCC AAGTATGAAT 1560
55      CTCACCTTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTAC 1620
GAGCAGACTC ACAGAGCCCC TCCCACITTG ATCTCAGCA GCCGTCAGTC TTGGAAGAG 1680
AAGAGGTCTA GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG 1740
GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT CGGCCAGTCA GACGATCTCA 1800
TTCAACACCA TCATGATAC CATCATATTC TCCATCATCA CCACCACTCA AACCACTATC 1860
60      CTCACAGTCA CAGCCAGGCG TACTCTGGGG AGGAGCTGAA AGATGCCGGC GTGCCACTT 1920
TGGCTCGGAT GGTGATAATG GGTGATGGCC TGCACATTT CAGCGATGGC CTAGCAATG 1980
GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTAAAGTAC TTCGTGTGCT GTGTTCTGTC 2040
ATGAGTTGCC TCATGAATTA GGTGACTTGG CTGTTCTACT AAGGCTGGC ATGACCGTTA 2100
AGCAGGCTGT CCTTATAAT GCATTGTGAG CCATGCTGGC GTATCTTGGG ATGGCAACAG 2160
65      GAATTTCTAT TGGTCAATAT GCTGAAATG TTTCTATGAG GATATTGCA CTACTGCTG 2220
GCTTATTCAT GTATGTTGCT CTGGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA 2280
GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTACA GAATGCTGGG ATGCTTTTGG 2340
GTTTGGAAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATCGTGTTC CGTATAAATT 2400
70      TTTAGTTAAG GTTAAATGCT TAGAGTAGCT TAAAAAGTTG TCATAGTTTC AGTAGGTCAT 2460
AGGAGATGTA GTTTGTATGC TGTACTATGC AGGTTTAAA GTTAGTGGGT TTTGTGATT 2520
TTGTATTGAA TATGCTGCTC TGTACAAAG TCAGTTAAAG GTACGTTTAA ATATTTAAGT 2580
TATTCATCTC TGGAGATAAA ATCTGATGT GCAATTCACC GGTATTACCA GTTTATTATG 2640
TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTCAG GGAATAATGT CTTTAAATGT 2700
75      TTTCAAGAA CTACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACGCTGCG 2760
TGTTTAGGAA TAAGAAATGT CATGAAGCCT AAATACCAA GAAGCTTAT ACTGAATTTA 2820
AGCAAGAAAA TAAGAGAA AAGAGAGAA TCTGAGAAAT GGGGAGGCAT AGATTCTTAT 2880
AAAAATCACA AATTTTGTG TAAATTAGAG GGGAGAAAT TAGAATTAAG TATRAAAGG 2940
CAGATTAGT ATAGAGTACA TTCAATAAAC ATTTTGTCA GGATATTTC CGTAAATAAC 3000
GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTAAAC TTGTGATAAT ACAGAAATCT 3060
80      AAATATATTT AATGAATCCA AGCAATATAC ACTTGACCAA GAAATTTGAA TTTCAAAATG 3120
TTCTGTGCGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGATACAC CAGACTGGGT 3180
TATTTGCCAG TTATATATCA CCAAAAGCTG TATGACTGGA TGTTCTGGTT ACCTGGTTTA 3240
CAAAATATTC AGAGTAGTAA AACTTTGATA TATATAGGGA TATTAAACT ACACTAAGTA 3300
TCATTTGATT CGATTAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATGTT 3360

```

GAGCAATTGT CTTTATATAC GGTACTGTAG CCATACTAGE CCTGTCTGTG GCATTCTCTA 3420
GATGTTTCTT TTTTACACAA TAAATTCCTT ATATCAGCCT G

5

A161 PROTEIN SEQUENCE

Gene name: LIV-1 protein, estrogen regulated
Unigene number: Hs.79136
Protein Accession #: NP_036451
Signal sequence: 1-21
Pfam domain: Zip[591-743]
Transmembrane domains: 330-346, 352-368, 427-444, 663-679, 688-703, 730-745
Cellular Localization: plasma membrane

10

15

1	11	21	31	41	51	
MARKLSVILI	LTFALSVTNP	LHELKAAAFP	QTTEKISPNW	ESGINVDLAI	STRQYHLQOL	60
FYRYGENNSL	SVEGFRKLLQ	NIGIDKIKRI	HIHHDHDSHS	DREHHSDBER	HSDHEHRSDE	120
EHSDHDSHS	HHMHAASGN	KRKALCPDHD	SDSSGKDPEN	SQSGKAHRPE	HASGRNRVKD	180
SVSASEVTST	VYNTVSEGT	FLETIETPRP	GKLFPPKDVSS	STPPSVTSKS	RVSRLAGRKT	240
NEGVSEPRKQ	FMYSRNTNEN	POECFNASKL	LTSHGMIQV	PLNATEFNYL	CPATINQIDA	300
RSCLIHSTSEK	KAEIPFKTYS	LQIANVGGFI	AISLIISFLSL	LGIVLVPIMN	RVFFKFLLSF	360
LVALAVGTLS	GDFAFLKLLPH	SHASHHSHS	HEEPAMEMKR	GPLFSLHSSQ	NIEBSAYFDS	420
TWKGLTALGG	LYFMFLVEHV	LFLIKQPKOK	XKKNQKPEP	DDVSEIKRQL	SKYESQLSTN	480
EEKVDTDRT	EGYLRADSQE	PSHFDSQQPA	VLEKEEVMIA	HAHPQEVYNE	YVPRGCKNKC	540
HSFHDTLGG	SDDLIHEHHD	YHEILHHEHH	QNHHPHSESQ	RYSREELKDA	GVATLAWMVI	600
MGGDLHNFSD	GLAIGAPFE	GLSSGLSTSV	AVFCHPELPE	LQDFAVLLKA	GMTVKQAVLY	660
NALSAMLAIV	GMATGIFIGH	YAENVSMWIF	ALTAGLFMYV	ALVDMVPEML	HNDASDHGCS	720
RWGYFFLQNA	GMLLFGGIML	LISIFEHKIV	FRINF			

30

A162 DNA sequence

Gene name: bone morphogenetic protein receptor IB (ALK-6)
Unigene number: Hs.87223
Probeset Accession #: AA250737
Nucleic Acid Accession #: NM_001203
Coding sequence: 274-1782 (underlined sequences correspond to start and stop codons)

35

40

1	11	21	31	41	51	
CGCGGGGCGC	GGAGTCGGCG	GGGCGCTCGC	GGAGCGGGGC	AGTCGGGAGA	CGCGGGGCGCT	60
GAGGACGCGG	GAGCGGGGAG	CGCACGCGCG	GGGTGGAGTT	CAGCCTACTC	TTTCTTAGAT	120
GTGAAGAGAA	AGGAAGATCA	TTTCATGCCCT	TGTTGATAAA	GGTTCAGACT	TCCTGCTGAT	180
CATAACCATF	TGGCTCTGAG	CTATGACAA	AGAGGAACA	AAAAGTTAA	CTTACAAGCC	240
TGCCATAGT	GAGAAGCAAA	CTTCTTGAT	AACATGCTTT	TGCGAAGTGC	AGGAAATTTA	300
AATGTGGCA	CCAAGAAAGA	GGATGGTGAG	AGTACAGCCC	CCACCCCCCG	TCCAAAGGTC	360
TTGCGTTGTA	AATGCCACCA	CCATTGTCCA	GAAGACTCAG	TCAACAATAT	TTGCAGCACA	420
GAOAGATAT	GTTCACGAT	GATAGAAGAG	GATGACTCTG	GGTTGCCTGT	GGTCACTTCT	480
GGTTGCCTAG	CATAGAAGG	CTCAGATTT	CAGTGTGGG	ACACTCCCAT	TCCTCATCAA	540
AGAAATCA	TTGATGCTG	CACAGAAAGG	AACGAATGTA	ATAAGACCT	ACACCTTACA	600
CTGCTCCAT	TGAAAAACAG	AGATTTTGT	GATGGACCTA	TACACCCAG	GGCTTTACTT	660
ATATCTGTGA	CTGTCTGTAG	TTTCTCTCTG	GTCTTTATCA	TATTATTTCG	TTACTTCCGG	720
TATAAAGAAC	AAGAAACACG	ACCTCGATAC	AGCATTGGGT	TAGAACAGGA	TGAAACTTAC	780
ATTCCTCTCT	GAGAATCCCT	GAGAGACTTA	ATTGAGCAGT	CTCAGAGCTC	AGGAAGTGA	840
TCAGGCTTCC	CTCTGCTGGT	CCAAAGGACT	ATAGCTAAGC	AGATTTCAGT	GGTGAACAG	900
ATTGGAAG	GTGCTATGG	GGAGTTTGG	ATGGGAAAGT	GGCGTGGCGA	AAAGGTAGCT	960
GTGAAAGTGT	TCTTCACAC	AGAGGAAGCC	AGCTGGTTCA	GAGAGACAGA	AATATATCAG	1020
ACAGTGTGTA	TGAGGCATGA	AAACATTTTG	GGTTTCATTG	CTGCAGATAT	CAAAGGGACA	1080
GGGTCTGTGA	CCAGTTGTGA	CCTAATCACA	GACTATCATG	AAAATGGTTC	CCTTTATGAT	1140
TATCTGAAGT	CCACCACTCT	AGAAGCTAAA	TCAATGCTGA	AGTTAGCTTA	CTCTTCGTGC	1200
AGTGGCTTAT	GTCTTTTACA	CACAGAAATC	TTTAGTACTC	AAGGCAAAAC	AGCAATTGCC	1260
CATCGAGATC	TGAAAAGTAA	AAACATTTCT	GTGAAGAAAA	ATGGAACCTG	CTGTATTGCT	1320
GAOCTGGGCC	TGGCTGTATA	ATTATATTAGT	GATACAAATG	AAGTTGACAT	ACCACCTAAC	1380
ACTCGAGTGT	AGCCCAAACG	CTATATGCCCT	CCAGAAGTGT	TGGACGAGAG	CTTGAACAGA	1440
AATCACTTCC	GCTCTTACAT	CATGGCTGAC	ATGTATAGTT	TGGCCTCAT	CCTTTGGGAG	1500
GTGTCTAGGA	GATGTGTATC	AGGAGGTATA	GTGGAGAAAT	ACCAGCTTCC	TTATCATGAC	1560
CTAGTGCCCA	GTGACCCCTC	TTATGAGGAC	ATGAGGGAGA	TTGTGTGCAT	CAAGAAGTTA	1620
CGCCCCCTAT	TCCCAAACCG	GTGGAGCAGT	GATGAGTGTG	TAGGCGAGAT	GGGAAACTC	1680
ATGACAGAA	CTGCGGCTCA	CAATCTCTCA	TCAAGGCTGA	CAGCCCTGCG	GGTTAAGAAA	1740
ACACTTGCCA	AAATGTCTCA	GTCCCAAGAC	ATTAACTCT	GATAGGAGAG	GAAAAGTAA	1800
CATCTCTGCA	GAAGGCCAAC	AGGTACTCTT	CTGTTTGTGG	GCAGAGCAAA	AGACATCAAA	1860
TAGCATCCA	CAGTACAAGC	CTTGAACATC	GTCTGCTTTC	CCAGTGGGTT	CAGACCTCAC	1920
CTTTACGGGA	GCGACCTGGG	CAAAGACAGA	GAAGCTCCCA	GAAGGAGAGA	TTGATCCGTA	1980
TCTGTITGTA	GGCGGAGAAA	CGTTTGGGTA	ACTTGTTCAA	GATATGATGC	AT	

75

A163 Protein sequence

Gene name: bone morphogenetic protein receptor IB (ALK-6)
Unigene number: Hs.72472 / Hs.87223
Probeset Accession #: AA250737 / U89326
Protein Accession #: NP_001194
Signal sequence: 1-13
Transmembrane domains: 128-144
PFAM domains: activin receptor [30-111], protein kinase [204-491]
Cellular Localization: plasma membrane

80

```

1      11      21      31      41      51
5      MLRSAGKLN VGTKKEDGES TAPTPRPKVL RCKCHHCPE DSVNNICSTD GYCFMIEED 60
      DSGLPVVTSS CLGLEGSDFQ CRDTPPEQR RSTECCTERN ECKNDLHPTL PFLKNRDFVD 120
      GPIHRRALLI SVTVCSLLLV LILFLCYERY KRQSTRPRYS IGLEQDETYI PPGESLRDLI 180
      EQSQSSGSGS GLPLLVQRTI AKQIQMVKQI GKGRYGEVVM GKWRGKAV KVPFTTRES 240
      WPRETEIYQT VLMRHNILG FIAADIKGTG SWTQLYLITD YHENGSLVDY LKSTILDAKS 300
10     MLKLAYSSVS GLCHLHTEIF STQKPAIAH RDLKSKNILV KNGTCCID LGLAVKFIID 360
      TNEVDIPPNT RVGTIKRYMFP EVLDESILNRN HFQSYIMADM YSFLILNEV ARRCVSGGIV 420
      EBYQLPYHDL VPSPSYEDM REIVCIKKLR PSFENRNSD ECLRQMGKLM TECWAHPAS 480
      RLTAIRVKKT LAKMSQDI KL

```

A164 DNA sequence

```

15     Gene name: ESTs
      Unigene number: Hs.157601
      Probeset Accession #: W07459
      Nucleic Acid Accession #: AC005383
20     Coding Sequence: 328-2751 (underlined sequences correspond to start and stop codons)

```

```

1      11      21      31      41      51
25     GACAGTGTTC GCGGCTGCAC GCTCGGAGG CTGGGTGACC CGGTAGAAAG TGAAGTACTT 60
      TTTTATTTCG AGACTGGGCG CGATGCCGCT TTAATAAACG CGAGGGGCTC TATGCACCTC 120
      OCTGGCGGTA GTTCCTCCGA CCTCAGCCGG GTCCGGTGCT GCGGCCCTCT CCCAGGAGAG 180
      ACHAAACAGT GTCCACAGTG GCAGCCCGCG CCGGGGCGCC CTTCTGTGA TCCGTAGCG 240
      CCCCCGCGCC CGAGCCCGCG CCGGGTCTGT GAGTAGAGCC GCGCGGCGAC CGAGCGCTGG 300
30     TCGCGCTCTT CCTTCGTTA TATCAACATG CCCCCTTTCC TGTTCCTGGA GCGCGTCTGT 360
      GTTTCCTGTG TTTCAGAGT GCCCCATCT CTCCCTCTCC AGGAAGTCCA TGTAAAGCAA 420
      GAAACCATCG GGAAGATTTC AGCTGCCAGC AAAATGATGT GGTGCTCGGC TGCAGTGGAC 480
      ATCATGTGTT TGTAGATGG GTCTAACAGC GTCCGGAAGG GGAGCTTTGA AAGGTCCAG 540
      CACTTTGCCA TCACAGTCTG TGACGGTCTG GACATCAGCC CGAGAGGGGT CAGAGTGGGA 600
35     GCATTCCAGT TCAGTCCAC TCCTCATCTG GAATTCCTCT TGGATTCAAT TTCACCCCAA 660
      CAGGAAGTGA AGCAAGAAAT CAAGAGGATG GTTTTCAAAG GAGGGCGCAC GGAGACGGA 720
      CTTGCTCTGA AATACCTCTT GCACAGAGGG TTGCTGGAG GCAGAAATGC TTCTGTGCCC 780
      CAGATCTCTA TCATCGATCG TGATGGGAAG TCCAGGGGG ATGTGGCACT GCCATCCAG 840
      CAGCTGAAGG AAAGGGGTGT CACTGTGTTT GCTGTGGGG TCAGGTTCCT CAGGTGGGAG 900
40     GAGCTGCATG CACTGCGCAG CGAGCCTAGA GGGCAGCACG TGCTGTGGC TGAGCAGGTG 960
      GAGGATGCCA CCAACGGCTT CTTCAGCAC CTCCATCTG CCGCATCTG CTCCAGCGCC 1020
      ACGCCAGACT GCGAGGTGGA GGCTCACCCC TGTGAGCACA GGACGCTGGA GATGGTCCGG 1080
      GAGTTGCTGT GCATGCCCC ATGCTGGAGA GATGCGGGC GGACCTTGC GGTGTGGCT 1140
      GCACACATGC CTTCTACAG CTGGAAGAGA GTGTTCCTAA CCAACCTGC CACTGTCTAC 1200
45     AGGACCACTT GCGCAGGCC CTGTGACTCG CAGCCTGCC AGAATGAGG CACATGTGTT 1260
      CCAGAAGGAT TCGAGGCTTC CAGTGCCTC TGCCGCTGCG CTTTGGAGG GGAGGCTAAC 1320
      TGTGCCCTGA AGCTGAGCTT GGAATGAGG GTGACCTCC TCTTCTGCT GACACGCTCT 1380
      GCGGGCAGCA CTCTGACGG CTTCCTGGGG GCCAAGTCT TCGTGAAGG GTTGTGGG 1440
      GCCGTGCTGA GCGAGGACTC TCGGGCCCGA GTGGGTGGG CCACATACAG CAGGGAGCTG 1500
50     CTGGTGGCGG TGCTGTGGG GGAGTACCAG GATGTGCTG ACCTGTGCTG GAGCCTCGAT 1560
      GGCATTCCCT TCCTGTGGG CCCCACCTGT ACGGGCAATG CTTGTGCGCA GCGGCGAGAG 1620
      CGTGGCTTCG GAGCGGCCAC CAGGACAGGC CAGGACCGGC CAGTAGAGT GGTGGTTTGG 1680
      CTTACTGAGT GCATGCTCGA GAGTAGGTT GCGGGCCAG CGCTCACGC AAGGGCGCGA 1740
      GAGCTGTCTC TGCTGGGTGT AGGCAGTGAG GCGGTGCGGG CAGAGCTGGA GGAGATCACA 1800
55     GGCAGCCCAA AGCATGTGAT GGTCTACTCG GATCCTCAG ATCTGTCAA CCAATCCCT 1860
      GAGCTGCAGG GGAAGCTGTG CAGCCGGCAG CGGCCAGGT GCCGACACA AGCCTGGAC 1920
      CTCGTCTTCA TGTGGACAC CTCTGCTCA GTAGGGCCCG AGAATTTGC TCAGATGCAG 1980
      AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT GAGGTGAACC CTGACGTGAC ACAGGTGCGC 2040
      CTGGTGGTGT ATGGCAGCCA GGTGCAGACT GCCTTGGGC TGGACACAA ACCACCCCG 2100
60     GCTGCGATGC TGCGGGCCAT TAGCCAGGCC CCTACCTAG GTGGGGTGGG CTGAGCCGGC 2160
      ACCGCCCTGC TGACATCTA TGACAAAGTG ATGACCGTCC AGAGGGGTGC CCGGCTGGT 2220
      GTCCCCAAGC CTGTGGTGGT GCTCACAGGC GGGAGAGCG CAGAGGATGC AGCGTTCTCT 2280
      GCCCAGAGC TGAGGAACAA TGGCATCTCT GTCTTGTG TGGGCTGGG GCCTGTCTTA 2340
      AGTAGGGTTC TGCGGAGGCT TGCAGGTCC CCGGATTCCT TGATCCAGT GGCAGCTTAC 2400
65     GCGGACCTGC GGTACACCA GACGTGCTC ATTGAGTGGC TGTGTGGAGA AGCCAAGCAG 2460
      CCAGTCAACC TCTGCAAAAC CAGCCCGTGC ATGAATGAGG GCAGCTGCGT CCTGCAGAA 2520
      GGGAGCTACC GCTGCAAGTG TCGGATGCG TGGGAGGGCC CCACTGCGA GAACCGTGAG 2580
      TGGAGCTCTT GCTCTGTATG TGTGAGCCAG GGATGGATT TTAGAGAGCC CCTGAGGCAC 2640
      ATGGCTCCCG TGCAAGGAGG CAGCAGCGT ACCCTGCCA CCAACTACAG AGAAGGCCTG 2700
70     GGCAGTGAAT TGGTGCCTAC CTTCGTGAAT GTCTGTGCC CAGGTCTCTA GAATGTCTGC 2760
      TTCCGCGTAC TGGAGCAGCC ACTATTCTCA CTGAGGAGG AGGATGTCCC AACTGCAGCC 2820
      ATGCTGTCTA GAGACAAGAA AGCAGCTGAT GTCACTTACA AACGATGTT TGAAGAGTT 2880
      TTGATGTCTA GTTAAATACC CACTTCTGT ACCTGCTGTG CTTTGTGAG GCTATGTCT 2940
      CTGCCACCTT TCCCTGAGG ATAAACAAGG GGCTCTGAG ACTTAAATTT AGCGGCTTGA 3000
75     CGTTCCTTTG CACACATCTA ATGCTGCGCA GAATGTTGTT GACACAGTAA TGCCCGCAG 3060
      AGGCTTTTAC TTTGGACGCG TTTGGACGCG GAAGGCCAAG GCCTTTCAAG ATGGAAGCA 3120
      GCAAGCTTTT CACTTCCCCA GAGACATTCT GGATGCATT GCATTGAGTC TGAAGGGGG 3180
      CTTGAGGGAG GTTGTGTACT TCTTGGCGAC TGCTTTTGT GTGTGGAAGA GACTTGAAGA 3240
      GGTCTCAGAC TGAATGTGAC CAATTAACCA GCTTGGTTGA TGATGGGGA GGGGCTGAGT 3300
80     TGTGATGGG CCGAGTCTG GAGGGCCAG TAAATCGTT CTGAGTCTG AGCAGTGTCC 3360
      ACCTTGAAGG TCTTC

```

A165 Protein sequence

Gene name: ESTs
 Unigene number: Hs.157601
 Protein Accession #: none found
 Signal sequence: 1-17
 Transmembrane domains: none found
 VGM domains: 49-223; 341-518; 529-706
 EGF domains: 298-333; 715-748
 Cellular Localization: secreted

```

1      11      21      31      41      51
|      |      |      |      |      |
15  MPPFLLEAV CVFLFSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLLDGSN 60
    SVGKGSFERS KHFALTVC DG LDISPERVRV GAFQFSSTPH LEFPIDSPST QQEVKARIKR 120
    MVEKGSRTET ELAKYLLHR GLPGGRNASV PQILIIIVTDG KSQGDVALPS KQLKERGVTV 180
    FAVGVRFPPR BELHALASEP RGQSVLLAEQ VEDATNGLFS TLSSSAICSS ATPDCRVEAH 240
    PCEHRTLELV REFAGNAPCW RGSRRTLAVL AAHCIFYSNK RVFLTHPATC YRTTCPGPCD 300
    SQPCQNGGTC VPEGLDGYQC LCPLAFGGEA NCALKLSLEB RVDLLFLDLS SAGTTLDGFL 360
    RAKVFVKRFV RAVLSSEDSRA RVGVATYSRE LLVAVFVGEY QDVPDLVWSL DGLFFRGGPT 420
    LTGSAALRQAA ERGFGSATRT GQDRPRRVVV LLTSEHSEDE VAGPARHARA RELLLLGVS 480
    EAVRAELSEI TGSPIHVMVY SDPQDLFNQI PELQKGLCSR QEPGCRTOAL DLVEMLDISA 540
    SVGPEHFAQM QSPVRSCALQ FEVNFEDVTQV GLVVYGSQVQ TAFGLDTKPT RAAMLRAISQ 600
    APYLGSVGSA GTALLHIYDK VMTVQRGARP GVEKAVVVLT GGRGASDAV PAQKLRRNGI 660
    SVLVVGVGVF LSEGLRRLAG PRSLIHVAA YADLRVHQDV LIENLCGEAK QFVNLCKPSP 720
    CMNNGSCVLQ NGSVRCKCRD GWSGPHCENR EWSSCSVCVS QGWILETPLR HMAFVQEGSS 780
    RTPPSNYREG LGTEMVPTFW NVCAQGP
  
```

A166 DNA sequence

Gene name: Homo sapiens beta-1 adrenergic receptor mRNA, 3' UTR
 Unigene number: Hs.37744
 Probeset Accession #: AA011176
 Nucleic Acid Accession #: AF272890
 Coding Sequence: 87-1520 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
|      |      |      |      |      |
40  TGCTACCCGC GCCCGGGCTT CTGGGGTGTG CCCCAACCAC GCGCCAGCCC TGCCACACCC 60
    CCCGCCCTCG GCTTCGCGAG CTCGGCATGG GCGCGGGGGT GTCCTCTCTG GCGCCCTCGC 120
    AGCCCGGTAA CCTGTGCTCG GCGGCACCGC TCCCCGACGG CCGCGCCACG GCGCGCGCGC 180
    TGCTGGTGCC CGCGTGCCTG CCGCGCTCGT TGCTGCTTCC GCGCAGCGAA AGCCCGGAGC 240
    CGCTGTCTCA GAGTGGACA GCGGCGATGG GTCTGCTGAT GCGCTCTATC GTGCTGCTCA 300
    TCGTGGCGGG CAATGTGCTG GTGATCGTGG CCATCGCCAA GACGCGCGGG CTGCAGACGC 360
    TCACCAACCT CTTCATCATG TCCCTGGGCA GCGCGGACCT GGTCTATGGG CTGCTGGTGG 420
    TGCCGTTCGG GCGCCACATC GTGGTGTGGG GCGCGCTTCC GTACGCTTCC TTCTTCTGGG 480
    AGCTGTGGAC CTCACTGGAC GTGCTGTGGG TGACGCGCAG CATCGAGACC CTGTGTGTCA 540
    TTGCCCTGGA CGCTCACTTC GCCATCACCT GCGCCTTCCG CTACGAGAGC CTGCTGACGC 600
    GCGCGCGGGC GCGGGGCTTC GTGTGCAAGG TGTGGGCCAT CTGCGCCCTG GTGTCTCTCC 660
    TGCCCATCTT CATGCACTGG TGGCGGGCGG ABAGCGAAGA GCGCGCCCGC TGCTACACAC 720
    ACCCCAAAGT GCGCGACTTC GTCAACCAAC GCGCCTAAGC CATCGCCTCG TCCGTAGTCT 780
    CCTTCTAAGT GCGCCTGTGC ATCATGGCTT TCGTGTACCT GCGGGTGTTC GCGGAGGCCC 840
    AGAAGCAGGT GAAGAAGATC GACAGCTGGC AGCGCGGTTT CCGCGCGGGC CCAGCGCGGC 900
    CGCCCTCGCC CTCGCCCTCG CCGGTCCCGG CGCCCGCGCC GCGCGCGGA CCCCGCGCC 960
    CGCGCGCGCG CGCGCGCACG GCCCGCTGGC CCAAGCGGGC TGCGGTAAAG CGCGCGCCCT 1020
    CGCGCGCTGT GCGCCTACGC GAGCAGAGAG CGCTCAAGAC GCTGGGATC ATCATGGGCG 1080
    TCTTCACGCT CAGCTGGCTG CCTTCTTCTC TGGCCAAAGT GGTGAAGGCC TTCCACCGCG 1140
    AGCTGTGTCC CGACCGCTTC TTGCTCTTCT TCAACTGGCT GGGCTACGCC AACTCGGCTG 1200
    TCAACCTCAT CATCTACTGC CGCAGCCCGG ACTTCCGCAA GCGCTTCCAG GGAAGTCTCT 1260
    GCTGCGCGCG CAGGGCTGCC CGCGCGCGCC ACGGACCCA CGGAGACCGG CGCGCGCGCT 1320
    CGGGCTGTCT GCGCGCGGCC GAGCCCCCGC CATGCGCGCG GCGCGCTTCC GAGCAGCAGC 1380
    ACGACGATGT GTGCGGGGCC ACGCGCGCGG CGCGCGCTGT GGAGCCCTGG GCGCGCTGCA 1440
    ACGCGCGGGC GCGCGCGGAC AGCACTCGA GCGTGAACGA GCGGTGCGCG CCGCGCTTGG 1500
    CCTCGGAATC CAAGGTGTAG GCGCGCGCGC GGGCGCGGGA CTCGCGGCAC GGCTTCCCAG 1560
    GGGAACGAGG AGATCTGTGT TTACTTAAGA CCGATAGCAG GTGAACCTGA AGCCACAAAT 1620
    CCTGTCTGTA ATCATCGGAG GCAAAGAGAA AAGCCACGGA CGTTGCACA AAAAGGAAAG 1680
    TTTGGGAAGG GATGGGAGAG TGGCTTGTCT ATGTCTCTTG TTG
  
```

A167 Protein sequence

Gene name: Homo sapiens beta-1 adrenergic receptor mRNA, 3' UTR
 Unigene number: Hs.37744
 Protein Accession #: AA011176
 Signal sequence: none found
 Transmembrane domains: 62-84, 95-117, 135-157, 177-198, 226-248
 Pfam domain: 7tm_1 [75-377]
 Cellular Localization: plasma membrane

```

1      11      21      31      41      51
|      |      |      |      |      |
80  MGAGVLVIGA SEPCNLSSAA PLEDGAATAA RLLVPASPPA SLLPPASESP EPLSQQWTAG 60
    MGLMALIVLI LIVAGNVLVI VAIKTERLQ TLTNLFIMSL ASADLVMSGL VVPPGATIVV 120
  
```

WCRWEVGSFF CELWTSVDVL CVTASIEILC VIALDRYLAI TSPFRYQSLI TRARARGLVC 180
 TVWAIASLVV FLFILMHMWR AESDEARRCY NDPKCCDFVT NRAYAIASSV VSFYVPLCIM 240
 AFVYLRFVFE AQKQVKIKDS CERRFLGGPA RPPSPSPSPV PAPAAPPSPV RPAAAAATAP 300
 LANGRACKRR PRLVALREBQ KALKTLGIIM GVFTLCWLEF FLANVVKAFH RELVFDRLFV 360
 FFWNLGYANS AFNP1IYCRS PDFKAPQGL LCCARRAARR RHATHGDRPR ASGCLARPGP 420
 PPSPGAASDD DDDDDVVGATP PARLLEPWAG CNGGAADSD SSDLPEPCRPV FASESKV

A168 DNA sequence

Gene name: CEGP1
 Unigene number: Hs.222399
 Probeset Accession #: AA256485
 Nucleic Acid Accession #: AJ400877
 Coding sequence: 81-3080 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 GGCGTCCGCG CACACCTCCC CGCGCCGCGC CGGCCACTGC CGCACTCCG CCGCCTCTGC 60
 CCGCAACCGC TGAGCCATCC ATGGGGGTGG CGGGCCGCAA CCGTCCCGGG GCGGCTTGGG 120
 CGGTGCTGCT GCTGCTGCTG CTGTGCGCGC CACTGTGCTG CTGCGCGGGG GCCCTCCCGC 180
 CGGTCCGGG CGGTGCCGCG GGGCCGCGAG AGGATGTAGA TGAGTGTGCC CAAGGGCTAG 240
 ATGACTGCCA TGCCGACGCC CTGTGTGAGA ACACACCCAC CTCCTACAAG TGCTCTGCCA 300
 AGCCTGGCTA CCAAGGGGAA GGCAGGCACT GTGAGGACAT CGATGAATGT GGAATAGAGC 360
 TCAATGGAGG CTGTGTCAT GACTGTTTGA ATATTCCAGG CAATTATCCT TGCACCTGTT 420
 TTGATGGCTT CATGTGGCT CATGACGGTC ATAATTGTCT TGATGTGGAC GAGTGCCTGG 480
 AGAACAAATG CGGCTGCCAG CATACCTGTG TCAACGTCAT GGGGAGCTAT GAGTGTGCT 540
 CGAAGGAGGG GTTTTCTCTG AGTGACAAATC AGCACACCTC CATTCACCGC TCGAAGAGGG 600
 GCGTGGAGCT CATGAATAAG GATCACGGCT GTAGTCACT CTGCAAGGAG GCCCCAGGGG 660
 GCAGCGTCCG CTGTGAGTCC AGGCTGTGTT TTGAGCTGGC CAAGAACCAG AGAGACTGCA 720
 TCTTGACCTG TAACCATGGG AACGGTGGGT GCCAGCACTC CTGTGACGAT ACAGCCGATG 780
 GCCCAGAGTG CAGCTGCCAT CCACAGTACA AGATGCACAC AGATGGGAGG AGCTGCCCTG 840
 AGCGAGAGGA CACTGTCTTG GAGGTGACAG AGAGCAACAC CACATCAGTG GTGATGGGG 900
 ATAAACGGGT GAAACGGGCG CTGCTCATGG AAACCTGTGC TGTCAACAAAT GGAGGCTGTG 960
 ACCGACCTG TAAGGATACT TCGACAGGTG TCCACTGCAG TTGTCTCTGT GGATTCACTC 1020
 TCCAGTGGGA TGGGAAGACA TGTAAAGATA TTGATGAGTG CCGAGCCCGC AATGAGGTT 1080
 GTGATCATTT CTGCAAAAC ATCTGTGGCA GTTTTGACTG CGGCTGCAGG AAGGATTTA 1140
 AATTATTAC AGATGAGAG TCTTGCCAG ATGTGGATGA GTGCTCTTGG GATAGGACCT 1200
 TGACACACAG CTGCATCAAC CACCTTGCA CATTGTCTG TGTGTGCAAC CGAGGGTACA 1260
 CCGCTGATCT GTGTCAACAC TGTGGAGACA CCAATGAGTG CAGCATCAAC AACGGAGGCT 1320
 GTGAGCAGGT CTGTGTGAAC ACAGTGGGCA GCTATGAATG CCAAGTCCAC CCTGGGTACA 1380
 AGCTCCACTG GAATAAAAA GACTGTGTGG AAGTGAAGGG GCTCTGCGCC ACAAGTGTGT 1440
 CACCCGCTGT GTCCCTGCACT TGCGGTAAAG TGGGTGGAGG AGACGGGTGC TTCTCTAGAT 1500
 GTCACTCTGG CATTCACCTC TCTTCAGATG TCAACCATC CAGGACAGT GTAACCTTTA 1560
 AGCTAAATGA AGCGAGGTGT AGTTTGAATA ATGCTGAGCT GTTTCCCBAG GGTCTGCGAC 1620
 CAGCACTACC AGAGAGACAC AGCTCAGTAA AAGAGAGCTT CCGCTACGTA AACCTTACAT 1680
 GCAGCTCTGG CAAGCAGTC CCAGGAGGCC CTGGCCGACC AAGCACCCCT AAGGAAATGT 1740
 TTATCACTGT TGAATTTGAG CTGAAACTA ACCAAAGGA GGTGACAGCT TCTGTGACAT 1800
 TGAGTGTGAT CGTAAAGCGA ACCGAGAGGC GGCTCCGTAA AGCCATCCGC ACCTCTAGAA 1860
 AGCGCGTCCA AGCGAGCGAG TTTCACCTCC AGCTCTCAGG CATGAACCTC GACGTGGCTA 1920
 AAAAGCCTCC CAGAACATCT GAACGCCAGG CAGAGTCTGG TGGAGTGGGC CAGGCTCATG 1980
 CAGAAACCA ATGTGTCACT TGCAGGCGTG GGACCTATTA TGATGGAGCA CGAGAACGCT 2040
 GCATTTTATG TCACAAATGA ACCTTCCAAA ATGAGGAAGG ACAATGACT TGTGAACCAT 2100
 GCCCAGAGAC AGGAAATCTT GGGGCCCCGA AGACCCAGA AGCTTGGAAAT ATGTCTGAAT 2160
 GTGGAGGTCT GTGTCAACCT GGTGAATATT CTGCAGATGG CTTTGCACCT TGCCAGCTCT 2220
 GTGCCCTGGG CAGCTTCCAG CTTGAAGCTG GTCGAACTTC CTGCTTCCCC TGTGGAGGAG 2280
 GCTTGGCCAC CAACATCAG GGAGCTACTT CCTTTCAGGA CTGTGAACCC AGAGTTCAT 2340
 GTTCACCTGG ACATTTCTAC AACACACCA CTCACCGATG TATTCTTGGC CCAGTGGGAA 2400
 CATACAGGCG TGAATTTGGA AAAAATAAT GTTTTCTCTG CCCAGGAAAT ACTACGACTG 2460
 ACTTGTATGG CTCCACAAAC ATAACCCAGT GTAAAAACAG AAGATGTGGA GGGGAGCTGG 2520
 GAGATTTTAC TGGGTACATT GAATCCCAAA ACTACCCAGG CAATTACCCA GCCAACACCG 2580
 AGTGAAGCTG GACCATCAAC CCACCCGCCA AGCGCCGCAT CCGATCGTG GTCCCTGAGA 2640
 TCTTCTGCC CATAGAGGAC GACTGTGGGG ACTATCTGGT GATGCGGAAA ACCTCTTCAT 2700
 CCAATCTCTT GACAACATAT GAAACCTGCC AGACCTACGA AGCCGCCATC GCCTTCAAGT 2760
 CCAGTCAAA GAAGCTGTGG ATTCAGTTCA AGTCCAATGA AGGGAACAGC GCTAGAGGGT 2820
 TCCAGTCCCC ATACGTGACA TATGATGAGG ACTACAGGA ACTCATTGAA GACATAGTTC 2880
 GAGATGGCAG GCTCTATGCA TCTGAGAAC ATCAGGAAAT ACTTAGGAT AAGAACTTTA 2940
 TCAAGGCTCT GTTTGATGTC TCGGCCATC CCCAGAACTA TTTCAAGTAC ACAGCCGAGG 3000
 AGTCCCGAGA GATGTTTCCA AGATCGTTCA TCGGATGCT ACCTTCCAAA GTGTCCAGGT 3060
 TTTTGAAGAC TTACAAATGA CTCAGCCAC GTGCCACTCA ATACAAATGT TCTGCTATAG 3120
 GGTGTGGTGG ACAGAGCTGT CTTCTTCTG CATGTGAGCA CAGTCCGGTA TTGCTGCTC 3180
 CGTATCAGT GACTCATTAG AGTTCAATT TTATAGATA TACAGATATT TTGGTAAATT 3240
 GAACCTGGTT TTTCTTTCCC AGCATCGTGG ATGTAGACTG AGAATGGCTT TGAGTGGCAT 3300
 CAGCTTCTCA CTGCTTGGG CGGATGTCTT GGATAGATCA CGGGCTGGCT GAGCTGGACT 3360
 TTGCTCAGC TAGGTGAGAC TCACTGTGCC TTCTGGGGTC TTAATCTCC TCAAGGAGTC 3420
 TGTAGTGGAA AGGAGGCCAC AGAATAAGCT GCTTATCTG AAACCTCAGC TTCTCTAGC 3480
 CCGGCCCTCT CTAAGGGAGC CCTCTGCAT CTGTGTCAGG CTCTGACCAG GCAGAACAGG 3540
 CAAGAGGGGA GGAAGGAGA CCCCTGCAGG CTCCTCCAC CCACTTGAG ACCTGGGAGG 3600
 ACTCAGTTTC TCCACAGCCT TCTCCAGCCT GTGTGATACA AGTTTGTATC CAGGAAGTTG 3660
 AGTTCTAAGC AGTGTCTGAG AAAAAAAGAA GCAGAAAGRA TTAGAAATAA ATAAAAACTA 3720
 AGCACTTCTG GAGACAT

A169 Protein sequence

Gene name: CEGP1
 Unigene number: Hs.222399

Probeset Accession #: AA256485
 Protein Accession #: CAB92285
 Signal sequence: 1-31
 Transmembrane domains: none
 PFAM domains: EGF-like_domains [49-84,132-167,177-213,286-321,407-442] CUE_domain [809-918]
 Cellular Localization: may be secreted

1 11 21 31 41 51
 10 MGVAQRNRPG AAVAVILLLL LEPPLLLLAG AVPPGRGAA GPQEDVDECA QGLDDCHADA 60
 LCQNPITSYK CSCKPGYQGE GRQCEDIDEC GNEINNGCVH DCLNIPGNYR CTCFDGFMIA 120
 HDGHNCLDVD ECLNNNGGCG HTCVNVMGSY ECCCKEGFPL SDNQHTCIHR SEBGLSCMNK 180
 DHGCSHICKE APRGSVACEC RPFELAKNQ RDCILTCNHG NGGCQHSRDD TADGPECSC 240
 15 PQYKMETDGR SCLEREDTVL EVTESNTTSV VDGDKRVKRR LLMSTCAVNN GGCDRTCKDT 300
 STGVHCSCPV GFTLQLDGKT CKDIDECQIR NGGCDHFCKN IVGSPDCGCK KGFKLITDEK 360
 SCQDVDECSL DRTCDHSCIN HPGTACACN RGYTLYGFTH CGDTNECSIN NGGCOQVCVN 420
 TVGSYECQCH FQYKLNHWKK DCVEVKGLLP TSVPFRVSLH CGKSGGGGDC FLRCHSGIHL 480
 20 SSDVTITRTS VTFKLNEGKC SLKNALFFP GLRPALPEKH SSVKESFRYV NLTCSSGKQV 540
 PGAPGRPSTP KEMFITVEFE LETNQKEVTA SCDSLSCIVKR TEKRLKRAIR TLRKAVHREB 600
 FHLQLSGMNL DVAKKPPTS ERQAESCGVG OGHAEQCQVS CRAGTYDGA RERCILCPNG 660
 TFGNEBQMT CEFPCPRPGNS GALKTPERAWN MSECGLCQF GEYSADGFAP CQLCALGTFQ 720
 PEAGRTSCFP CGGGLATKHQ GATSFQDCST RVQCSPGHFY NTHTRCIRC PVGTYPPEFG 780
 25 KNNCVSCFBN TTTDFDGSIN ITQCKNRRCG GELGDFGYI ESENYPGNYV ANTECTWTIN 840
 PPPKRRLIV VPRIFLPIED DCGDYLVMRK TSSNSVTTY ETCQTYERPI APTSRSKKLM 900
 IQPKSNEGNS ARGFQVFPVT YDEYQELIS DIVRDGRLYA SENEQELIKD KKLKALPDV 960
 LAHPQNYFKY TAQESREMPF RSFTIRLLRSK VSRFLRPYK

A170 DNA sequence

Gene name: DEME-6 protein (KIAA0452)
 Unigene number: Hs.125783
 Probeset Accession #: AL039402
 Nucleic Acid Accession #: AF007170
 Coding sequence: 1-1725 (underlined sequences correspond to stop codon)

1 11 21 31 41 51
 35 AAGGAGGCGG CCTCCGGGAA AAGCGACCGC AGGACTCCTG AGAGCAGCCT CCATGAGGCC 60
 CTGGACGAGT GCATGACCGC CCTGGACCTC TTCTCTACCA ACCAGTTCTC AGAAGCACTC 120
 40 AGCTACCTCA AGCCGAGAAC CAAGGAAAGC ATGTACCACT CACTGACATA TGCCACCATC 180
 CTGGAGATGC AGGCCATGAT GACCTTTGAC CCTCAGGACA TCTCTGCTGC CGGCACATG 240
 ATGAGAGAGG CACAGATGCT GTGTCAAGAG CACCGGAGGA AGTCTTCTGT AACAGATTCT 300
 TTCTGCGAGC TGCTGAACCG CCCCACCTCG GCGCAATTC A CTGAAGAAGA AATCCACGCT 360
 GAGGTCTGCT ATGCAGAGTG CTTGCTGCG AGAGCAGGCC TGACCTTCCT GCAGGACGAG 420
 45 AACATGCTGA GCTTCATCAA AGGCGGCATC AAAGTTGAA ACAGCTACCA GACCTACAG 480
 GAGCTGGACA GCTTTGTTCA GTCTCTACAA TACTGCAAGG GTGAGAACCA CCGGCACTTT 540
 GAAGGAGGAG TGAAGCTTGG TGTAGGGGCG TTCAACCTGA CACTGTCCAT GCTTCTACT 600
 AGGATCTCGA GGTCTGTTGA GTTTGTGGGG TTTTCAGGAA ACAGGACTA TGGGCTGCTG 660
 CAGCTGAGG AGGCGACAGC TCCGCTCTG TGCTCTGTGT CATGCTCCTG 720
 50 CTGTCTACCC ACACCTTCCT CACCTTCGTG CTGGTACTG GGAACGTCIA CATCGAGGAG 780
 GCGGAGAACG TCTTGAAGCC CTACCTGAAC CGGTACCCCTA AGGCTGCCAT CTCTCTGTT 840
 TTTGAGGGA GATTTGAAGT CATTAAAGGC AACATTGATG CAGCCATCCG GCGTTTCGAG 900
 GAGTGTCTGT AGGCCACGCA GCACGTGAAG CAGTTTACCC ACATGTGCTA CTGGGAGCTG 960
 55 AATGTGTGCT TCACTTACAA GGGCCAGTGG AAGATGTCTT ACTTCTAGCG CGACCTGCTC 1020
 AGCAGGAGA ACTGCTGCTC CAAGGCCACC TACATTACCA TGAAGGCCCG CTACCTCAGC 1080
 ATGTTTGGGA ACGAGGACCA CAAGCCGTTT GGGGACGAG AGGTGGAATT ATTTGAGGCT 1140
 GTGCCAGGCC TGAAGCTCAA GATTGCTGGG AAATCTCTAC CCACAGAGAA GTTTGCCATC 1200
 CGGAGTCCC GCGCTACTT CTCTCTCAAC CCTATCTCGC TGCCAGTGCC TGCTCAGGAA 1260
 60 ATGATGTACA TCTGGAACGG CTACGCCGTG ATTGGGAGGC AGCCGAAACT CACGGATGGG 1320
 ATACTTGAGA TTATCACTAA GGTCTGAAGG ATGCTGGAGA AAGGCCGAGA GAACGAGTAC 1380
 TCAGTGGATG ACGAGTGCCT GTTGAATTTG TTGAAAGGCC TGTGTCTGAA ATAACCTGGG 1440
 CGTGTCCAGG AGGCCGAGGA GAATTTAGG AGCATCTCTG CCAATGAATA GAAGATTAAA 1500
 TATBACCACT ACTTGATCCC AAAGCCCTCT CTGGAGCTGG CCTGCTGCTT TATGGAGCAA 1560
 65 GACAGAAACG AAGAGGCCAT CAACTTTTGT GAATCTGCCA AGCAAACTA CAAGAATTAC 1620
 TCCATGGAGT CAAGGACACA CTTTCAATC CAGGCAGCCA CACTCCAAGC CAAGTCTTCC 1680
 CTAGAGAACA GCAGCAGATC CATGCTCTCA TCACTGCTCT TGTAGCTTTG TGCAGCAGTT 1740
 CCGGCTGGA AGACAGAGAC AGCTGGACAG AGCTCTGAA AACATTTCAA AATACCCCT 1800
 CCCCTGCCC TGCCCTGCCT TTGGGGTCCA CCGGCACCTC AGTTGGATGG CACAACATAG 1860
 70 TGTATCCGTG CAGAAGCCGA GCTGGCATT TCAACAGTGT AGCCAAGGCG CTTTGCCAG 1920
 GGCAGAGCAG GTGGAGCCCT CTGCTGCCC TATCACACAT ACGGGTACTT GCTTTTCACT 1980
 GTGATGTTTA AGAGAAATGA TGAACAGTTT ACATTTTCTT TAGAAATACA TTGATGGGAT 2040
 CACAGTTGGC TTTAAAAACC AACCACTCTG AAGTCTTTGT CTTCACCTAT 2100
 TATCATCTGG AGTGAATCTT CTTTATATGA TGATGCCAAA GGGCAAAATG CTTTTCAGAT 2160
 75 TCACAGATTG GTGACGGAG GTCTCTCAGA GGAACCTGAG AATGCTGGAG 2220
 AGAGGCTAAG CCTCAGGCTT CAATGCTTCT GGGGTGGGCT ATGAGGATGT ACACAGACAC 2280
 CCACCTACCT ACTACTCACA CTTCTTTTCA CTCTTTTGT AAATTTCCAA TTTAAAAATC 2340
 AAGCAGCTCT TTTTAGTGAG ATAAATCTG AGCTCTCTCT TAGAAAAATC AATCTCTACC 2400
 80 AGTAGAAAT GCCAGGCTT GATGGAAGAG CTGTGTAGCC CTTTCTATGC CAAAGCCAGG 2460
 AAATTTGGGG GGCAGGAGGA GGTCTCAGA ATCCAGCTCT TATCTTTGCT GTATGCCAAA 2520
 CTGAAACCACT TGGGAATAAT TTATGAAACA TAAAAATCTT CTGTACTTCA CTCCAGGTA 2580
 CATTTGCTTA CTGACAGCAT TTTTGTAAA ACTGTTATTC TTGAAAAA AAAAATAAAA 2640
 AA

A171 Protein sequence

Gene name: DME-6 protein (KIAA0452)
 Unigene number: Hs.125783
 Probeset Accession #: AL039402
 Protein Accession #: AAC39582
 Signal sequence: none
 Transmembrane domains: 210-226
 Cellular Localization: plasma membrane

1 11 21 31 41 51
 | | | | |
 MTALDLPLTN QFSEALSYLK PRTKESMYHS LTYATILEMQ AMMTFDPQDI LLAGNMMKEA 60
 QMLCQRERRK SSVIDSFSSS VNRPTLQGET EEEIHAEVCY AECLLQRAAL TFLQDENMVS 120
 FIKGSIKVRN SYQTYKELDS LVQSSQYCKG ENHPHFEGGV KLGVGAFNLT LSLMPLTRILE 180
 LLEFVGFSGN KDYGLLQLEB GASGHSFRSV LCVMLLLCYH TFLTFLVLGTG NVNIEBAEKL 240
 LKPYLNRYPK GAIFLFFAGR IEVTKGNIDA AIRRFEECC E AQQHVKQFHE MCYWELMNCF 300
 TYKQGWKMSY FYADLLSKEN CWSKATYIYM KAAYLSMFGK EDHKPFDDDE VELFRAVPGL 360
 KLKIAGKSLP TEKFAIRKSR RYFSSNFI SL PVPALSMYI WNGYAVIGKQ PKLTDGILEI 420
 ITKAEEMLEK GPENEYSVDD ECLVKLLKGL CLKYLGVRQE ABEENFRSISA NEKKIKYDHY 480
 LIPNALLELA LLLMEQDRNE EAIKLESAX QNYKNYSMES RTHFRIQAA LQAKSSLENS 540
 SRSMVSVSVL

A172 DNA sequence

Gene name: EST
 Unigene number: Hs.200102
 Probeset Accession #: AL117406
 Nucleic Acid Accession #: none found

Coding sequence: 1-4044 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 | | | | |
 ATGACTAGGA AGAGGACATA CTGGGTGCCC AACTCTTCG GTGGCCCTCG GAATCGTGGC 60
 ATCGACATAG GCGATGACAT GGTTTCAGGA CTATTATTATA AAACCTATAC TCTCCAAGAT 120
 GGCCCTCTGA GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGGCAGC TGTCCACCAG 180
 TGGGGGAAGT ATGATGCTGC CTGAGAAACC ATGATTCCTT TCCGTCCCAA GCCGAGGTTT 240
 CTGCCCCCCC AGCCCCCTGA CAATGCTGGC CTGTCTCCT ACCTCACCGT GTCATGGCTC 300
 ACCCCGCTCA TGATCCAAAG CTACGGAGT CGCTTAGATG AGAACACCAT CCTCCACTG 360
 TCAGTCCATG ATGCTCTAGA CAAAATGTC CAAAGGCTTC ACCGCTTTG GGAAGAAGAA 420
 GTCTCAGGSC GAGGGATTGA AAAAGCTTCA GTGCTTCTGG TGATGCTGAG GTTCCAGAGA 480
 ACAAGGTGTA TTTTCGATGC ACTTCTGGGC ATCTGCTTCT GCATTGCCAG TGTACTCGGG 540
 CCAATATTGA TTATACCAAA GATCCCTGAA TATTCAAGAG AGCAGTTGGG GAATGTGTGC 600
 CATGGAGTGG GACTCTGCTT TGCCCTTTT CTCTCCGAAT GTGTGAAGTC TCTGAGTTTC 660
 TCTCTCCATG GGCATATCAA CCAACGACCA GCCATCAGGT TCCGAGCAGT TGTTCCTCTC 720
 TTTGCTCTTG AGAAGCTCAT CCAATTTAAG TCTGTAATAC ACATCACCTC AGGAGAGGCC 780
 ATCAGCTTCT TCACCGGTGA TGTAAACTAC CTGTTTGAAG GGGTGTGCTA TGGACCCCTA 840
 GTACTGATCA GCTGCGCATC GCTGGTCATC TGACGCAATT CTCTCTACTT CATTATTGGA 900
 TACACTGATC TTATTGCCAT CTATGCTAT CTCCCTGTTT TCCACTGGC GGTATTCTAG 960
 ACAAGAATGG CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCAT 1020
 GTGACCAATG AAGTTCTCAC TTGCATTAG CTGATTAAAA TGTACACATG GGAGAAACCA 1080
 TTTGCAAAA TCATTGAAG TATGAAAGT CTGACTTCT GCTCCAAACC TGGTGATGGC 1140
 ATGGCTCTTA GCATGCTGGC CTCTCTGAAT CTCTCTCGGC TGTCAAGTGT CTTTGTGCTC 1200
 ATTCAGTCA AAGCTCTCAC GAATTCGAAG TCTGCAAGTA TGAGGTTCAG GAAGTTTTC 1260
 CTCCAGGAGA GCGCTGTTT CTATGTCCAG ACATTACAAG ACCCCAGCAA AGCTCTGGTC 1320
 TTTGAGGAGG CCACCTTGTC ATGCCAACAG ACCTGTCCCG GGATCGTCAA TGGGGCACTG 1380
 GAGCTGGAGA GGAACGGGCA TGCTTCTGAG GGGATGACCA GGCTTAGAGA TGCCCTCGGG 1440
 CAGAGGAGAG AAGGGAACAG CTTGGGCCCA GAGTTGCACA AGATCAACCT GGTGTGTCTC 1500
 AAGGGGATGA TGTTAGGGGT CTGCGGCAAC ACGGGGAGTG GTAAAGACAG CCTGTGTGCA 1560
 GGCATCTCGG AGGAGATGCA CTGTCTGAG GGTCTCGTGG GGGTGAGGG AAGCCTGGCC 1620
 TATGTCCCCC AGCAGGCCCTG GATCGTCAGC GGAACATCA GGGAGAACAT CCTCATGGGA 1680
 GGCCTCATATG ACAAGGCCCG ATACCTCCAG GTGCTCCACT GTGCTCCCT GAATCGGGAC 1740
 CTGGAACCTC TGCCCTTTGG AGACATGACA GAGATTGGAG AGCGGGGCCCT CAACCTCTCT 1800
 GGGGGGACAG AACAGAGGAT CAGCCTGGCC CGCGCCCTCT ATTCCACCG TCAGATCTAC 1860
 CTGCTGGACG ACCCCTGTCT TGCTGTGGAC GCCACGTTGG GGAAGCACAT TTTTGAGGAG 1920
 TGCAATTAGA AGACACTCAG GGGGAAGACG GTCTCTCTGG TGACCCACCA GCTGCACTAC 1980
 TTAGAATTTT GTGGCCAGAT CATTTTGTGG GAAATGGGA AAATCTGTGA AAATGGAACT 2040
 CACAGTGAAT TAATGCAGAA AAAGGGGAAA TATGCCCAAC TTATCCAGAA GATGACAAAG 2100
 GAAGCCACTT CGGACATGTT GCAGGACACA GCAAGATAG CAGAGAGGCC AAAGGTAGAA 2160
 AGTCAGGCTC TGGCCACCTC CTGGAAGAG TCTCTCAACG GAAATGCTGT GCGGAGCAT 2220
 CAGCTCACAC AGGAGGAGGA GATGGAAGAA GGCTCCTTGA GTTGGAGGGT CTACCAACCAC 2280
 TACATCCAGG CAGCTGGAGG TTACATGGTC TCTTGCAATA TTTTCTCTCT CGTGGTGTG 2340
 ATGCTCTCTT TAACGATCTT CAGCTTCTGG TGGCTGAGCT ACTGGTTGGA GCAGGGCTCG 2400
 GGGACCAATA GCAGCCGAGG GAGCAATGGA ACCATGGCAG ACCTGGGCAA CATTGAGAC 2460
 AATCTTCAAC TGTCTCTCTA CAGCTGGTG TACGGGCTCA ACGCCCTGCT CCTCATCTGT 2520
 GTGGGGGTCT GCTCTCTCAG GATTTTCAAC AAAGTCAACA GGAAGGCATC CAGGGCCCTG 2580
 CACAACAGC TCTTCAACAA GGTTTTCCGC TGCCCCATGA GTTCTCTTGA CACCATCCCA 2640
 ATAGGCCBGC TTTTGAACCTG CTTCGACGGG GACTTGAAGC AGCTGGACCA GCTCTTGGCC 2700
 ATCTTTTCAG AGCAATTTCT GGTCTGTGCC TTAATGGTGA TGGCCGCTCT GTTGATTGTC 2760
 AGTGTGCTGT TCCCATATAT CCGTTAATG GAGGCATATA TCATGTTTAT TTGCTTCATT 2820
 TATTATATGA TGTTCAGAAA GGCCATCGGT GTGTTCAAGA GACTGGAGAA CTATAGCCCG 2880
 TCTCCTTTAT TCTCCACAT CCTCAATCT CTGCAAGGCC TGAGCTCCAT CCATGCTCAT 2940
 GGAATAACCTG AAGACTTTCT CAGCCAGTTT AAGAGGCTGA CTGATGGCCA GAATAACTAC 3000
 CTGCTGTGTT TCTATCTTTC CACACGATGG ATGGCATTGA GGCTGGAGAT CATGAACCAAC 3060
 CTGTGACCT TGGCTGTGTC CCGTTCGCTG GCTTTTGGCA TTCTCTCCAC CCGCTACTCC 3120

5	TTTAAAGTCA TGGCTGTCAA CATCGTGCTG CAGCTGGCGT CCAGCTTCCA GGCCACTGCC 3180	
	CGGATGTGCT TGGAGACAGA GGCACAGTTC ACGGCTGTAG AGAGGATACT GCAGTACATG 3240	
	AAGATGTGTT TCTCGGAAGC TCCTTTACAC ATGGAAGGCA CAAGTTGTCC CCAGGGGTGG 3300	
	CCACAGCATG GGGAAATCAT ATTTCAGGAT TATCACATGA AATACAGAGA CAACACACCC 3360	
	ACCGTGTCTC ACGGCATCAA CCTGACCATC CGCGGCCACG AAGTGGTGGG CATCGTGGGA 3420	
	AGGACGGGCT CTGGGAAGTC CTCTTGGGC ATGGCTCTCT TCCGCTTGGT GGAGCCCATG 3480	
	GCAGGCCGGA TTCTCATTTGA CGGCGTGGAC ATTTGCAGCA TCGGCTTGGG GGAATTGCGG 3540	
10	TCCAAGCTCT CAGTGATCCC TCAAGATCCA GTGCTGCTCT CAGGAACCAT CAGATTCAAC 3600	
	CTAGATCOCT TTGACCGTCA CACTGACCAG CAGATCTGGG ATGCCCTTGA GAGGACATTC 3660	
	CTGACCAAGG CCATCTCAAA GTTCCCCAAA AAGCTGCATA CAGATGTGGT GGAACACGGT 3720	
	GGAACTTCTT CTGTGGGGGA GAGGCAGCTG CTCTGCATTG CCAGGGCTGT GCTTCGCAAC 3780	
	TCCAAGATCA TCCTTATCGA TGAAGCCACA GCCTCCATTG ACATGGAGAC AGACACCCCTG 3840	
	ATCCAGCGCA CAATCCGTGA AGCCTTCCAG GGCTGCACCG TGCTCGTCAT TGCCACCCGT 3900	
15	GTCAACCACTG TGCTGAAGTC TGACCAATC CTGGTATGG GCAATGGGAA GGTGGTAGAA 3960	
	TTTGATCGGC CGGAGGTACT GCGGAAGAA CCTGGGTCTT TGTTCGCAGC CCTCATGGCC 4020	
	ACAGCCACTT CTTCAGTGAG ATAAGGAGAT GTGGAGACTT CATGGAGGCT GGCAGCTGAG 4080	
	CTCAGAGGTT CACACAGGTG CAGCTTCGAG GCCCAGCTC TGCGACCTTC TTGTTTGGAG 4140	
	ATGAGAATT CTCTGGGAAG CAGGGGTAAA TGTAGGGGGG GTGGGGATTG CTGGATGGAA 4200	
20	ACCCCTGAAT AGGCTACTTG ATGGCTCTCA AGACCTTAGA ACCCCAGAAC CATCTAAGAC 4260	
	ATGGGAATT CA GTGATCATGT GGTTCCTCCT TTAACCTACA TGCTGAATAA TTTTATAATA 4320	
	AGGTAAAGC TTATAGTTTT CTGATCTGTG TTAGAAGTGY TGCAAAATGCT GTACTGACTT 4380	
	TGTAAATAT AAAACTAAGG AAAACTCAAA AAAAAAAAA AAAAAAA	
25	<u>A173 Protein sequence</u>	
	Gene name: EST	
	Unigene number: Hs.200102	
	Probeset Accession #: AL117406	
	Protein Accession #: none found	
30	Signal sequence: none found	
	Transmembrane domains: 169-185, 199-215, 275-291, 304-320, 387-403, 770-786, 829-845, 907-923, 927-943, 1018-1034	
	PPAM domains: ABC transporter [502-673], ABC membrane region [163-432, 771-1060]	
	ATP-binding domains [508-516, 1139-1147]	
35	Cellular Localization: plasma membrane	
	1 11 21 31 41 51	
	MTKRTYVWP NSSGGLVNRG IDIGDDMVSG LIYKTYTLQD GPWSQQRNRP EAPGRNAVPP 60	
40	WKKYDAALRT MIPFRPKPRF PAPQPLDNAG LPSYLTVSWL TPLMIQSLRS RLIDENTIPPL 120	
	SVHDASDKNV QRLHRLNREE VSRRGIEKAS VLLVMLRQPR TRLIFDALLG ICFCIASVLG 180	
	PILIIPKILE YSEBQLGNVV HGVGLCFALF LSECVKSLSF SSSWIINQRT AIRFRAVSS 240	
	FAPEKLIQPK SVIHITSGEA ISFFTGDVNY LFEQVCYGPL VLIITCASLVI CSISSYFIIG 300	
	YTAFLAILCY LLVFLAVFEM TRMAVKAGHH TSEVSDQIR VTSVVLATCIK LIKMYTNEKP 360	
45	FAKIEGMEB LTFCSKPGDG MAPSMLASLN LLRLSVFFVP LAVKGLTNSK SAVMEFKKFF 420	
	LQESPVEYVQ TLQDPKALV FEZATLSWQQ TPCGIVNGAL ELERNCHASE GMTPRDALG 480	
	PEEGNSLGP ELHKINLVVS KGMMLGVCGN TGGKKSLLS ALLEEMHLE GSVGVQSSLA 540	
	YVPQANTVS GNIRKNTLMG GAYDKARYLQ VLECCSLNRD LELLPPFGDMT EIGERGLNLS 600	
	GGQQRISLA RAVYSDRQIY LLDDPLSAVD AHVGEHIPSE CLIKTLRGKT VVLVTHQLQY 660	
50	LSEFQQLILL ENKICGENT HSELMQKKGK YAOIQRMHK EATSDMLQDT AKIAEKPKVE 720	
	SQALATSLER SLNGNAVPEH QLTQEEEMEE GSLSWRVYRH YIQAGGSYMV SCIIFFPVVL 780	
	IVFLTIFSPW WLSYWLEQGS GTNSRRSNG TMADLGNAD NFQLSFFYQLV YGLNALLLIC 840	
	VGVCSGIFT KVTREASTAL HNKLFNKVFR CPMSFFDTIP TGRLLNCFAG DLEQLDQLP 900	
	IFBEQFLVLS LMIIVALLIV SVLSPYILLM GAILMVICFI YMMFKKAGV VFKRLNYSR 960	
55	SPFLSHILNS LQGLSSIVY GKTEDFISQF KRLTDAQNNY LLLFLSSTRW MALRLIINTN 1020	
	LIVLAVALFV AFGISSTFYS FKVMVNIIVL QLASSPQATA RIGLETEAQF TAVERILQYM 1080	
	KMCVSEAPLE MBGTSCPQGW PQHGIIIPQD YHMKYRDNTF TVLHGINTLI RGHVVGIVG 1140	
	RTGGKBSLS MALFRIVFPM AGRILIDGVD ICSTGLEDLR SKLSVIPQDP VLSGTRIPN 1200	
	LDPFDRHTDQ QINDALERTF LTKAISFKPK KLHTDVVENG GNFVSGERQL LCIARAVLEN 1260	
60	SKILLIDEAT ASDIMEYDTL IQRTIREAFQ GCTVLVIAHR VITVLNCDHI LVMKNGKVE 1320	
	FDRPEVLRKK PGLSFAALMA TATSSLR	
65	<u>A174 DNA sequence</u>	
	Gene name: ESTs	
	Unigene number: Hs.128899	
	Probeset Accession #: AA983251	
	Nucleic Acid Accession #: AA983251	
	Coding sequence: 1-1749 (underlined sequences correspond to start and stop codons)	
70	1 11 21 31 41 51	
	<u>ATGCTGTCTG</u> <u>GCTTCTTGAT</u> <u>GAGTCCCACT</u> <u>ACCCAGCACA</u> <u>GAGCACAGTA</u> <u>CACTCCCGGA</u> 60	
	GGAAAGAAAC TTCCGTGGGA GGCCTCCATC GGTGCGCACA CCTCCCGAGG GCGAGGCGAG 120	
75	GACCGGGAGA GGGAGAGCCG GCCCGAGGCT GCCGGGCTCC TGTGGGACCG CGCTGCAGCC 180	
	GGGAGGCGCG AGAAGGGGAA CCGGGGCGAG CCGCCCGCCT GATCCGCGCC CAGCAGCAG 240	
	CCGCGGCGGC CGCCAGCTGG GCAGGCTCCC GGGACTGCGG CTGGGGGCGC GCAGGACCTT 300	
	CGCTCGGCTC CTGAGCTTTC CCGGGGAGG GTCCGGTTGC CAGTGAAACC TCCAGAGGCT 360	
	TCGGAGCAGC AGCCCGGGGG GCTTCTGAC TGCAATCCGA GATTTCCATC AGCGAGTGCA 420	
80	ACTCATAAGG CAGTCCCTAA GGGGACCGGG CCACCGGCTG AGGACGGGGA TGGCTTAGGA 480	
	GCTCCTGGAC CTAGGCCCCG GCGTCGTGCG CTCCTGGGCG TCGCGGCAGA GGGGAGTGCC 540	
	CTCGCGCGAA AGCGCGCGCG GACAGTCAGT GACGAGGCCC GGGGGTGGCC GGGGCCACGA 600	
	CTCTCTGGAG ACCGTCTGCG GCTCTCTGGA GACGCGCTGT CCGCGCCAG GGTGGTGCCA 660	
	TGTGGGCGBC TCGCGGCTCG TCGCTCTCCT CATCTGGAA CGCGGCTTCC CTCTGCGAGC 720	
	TGCTGCTGGC TGCCTGCTCG GCGGCGGGGG CGAGGGCCCA GCGGCGAGTA CTGCCACGGC 780	

5	TGGCTGGACG	CGCAGGGCGT	CTGGCGCATC	GGCTTCCAGT	GTCCCGAGCG	CTTCGACGGC	840
	GGCGACGCCA	CCATCTGCTG	CGGCAGCTGC	GCCTTGCCCT	ACTGCTGCTC	CAGCGCCBAG	900
	GGCGCGCTGG	ACCAGGGCGG	CTGCAGCAAT	GACCGCCAGC	AGGGCGCTGG	CGAGCCCTGGC	960
	CGGGCGGACA	AAGACGGGCC	CCGACGGCTC	GGCAGGGCTT	CATGTCTTAG	GGGTACCCAA	1020
	GGAGACGGGG	AGGGTGGCGC	CCCACCCGTG	AGGGCTGGCG	AGCGGTGCTC	CCCTGAAGGC	1080
	TCCCCGAAAG	GAAGGCAGCT	CCTCAGGGCT	TTCCCGGGGC	TGCTGCCCGG	TGCCAGACGC	1140
	CGCGGATTCC	CATCTTCTCC	ACCGGGCGGC	CCCTCPCCCC	TGCAGCGGCC	CGCTTGGCCC	1200
	ATCTACGTGC	CGTTCTCTAT	TGTTGGCTCC	GTGTTTGTCT	CTTTATCAT	CTTGGGGTCC	1260
10	CTGGTGGCAG	CCCTGTGCTG	CAGATGCTC	CGGCTAAGC	AGGATCCCA	GCAGAGCGGA	1320
	GCCCCAGGGG	GTAACCGCTT	GATGGAGACC	ATCCCCATGA	TCCCCAGTGC	CAGCACCTCC	1380
	CGGGGGTCTG	CTTCAGGCCA	GTCCAGCACA	GCTGCCAGTT	CCAGCTCCAG	CGCCAACTCC	1440
	GGGGCCCCGG	CGCCCCAAC	AAGGTACAG	ACCAACTGTT	GCTTGCCGGA	AGGGACCATG	1500
	AACAACGTGT	ATGTCAACAT	CGCCACGAAT	TTCTCTGTGC	TGAACGTCA	GCAGGCCACC	1560
15	CAGATTGTGC	CACATCAAGG	GCAGTATCTG	CATCCOCCAT	ACGTGGGGTA	CACGTGTCAG	1620
	CAGCACTCTG	TGCCCATGAC	AGCTGTGCCA	CTTTTCATGG	ACGGCTGCA	GCCTGGCTAC	1680
	AGGCAGATTG	AGTCCCCCTT	CCCTCACACC	AACAGTGAAC	AGAGATGTA	CCAGCTGGTG	1740
	ACTGTATAAC	TAACCTGTAC	TGGTGGGTTT	CTTTACTGAA	GGGAGACGAA	GGCAGGGGTG	1800
	GATTTCTGAG	GTGGAGAGTC	GCACATGTGC	GTGGTATTTA	TGGCACGACT	CTTTGGGATG	1860
20	GCCTTCTTTG	CCCCCAGACT	GTATGAAAAC	ATCTCGAAT	TAGCATTTCT	GGATATGTTT	1920
	CATCCAGGGT	ATCATTGATT	TATGATGGA	AACCGGCCCT	AGCTGGAGAT	GACTGTGATG	1980
	TTGCTGATGG	GGTATATAAC	AATGCTTGAG	TCCGAGTGC	CCCTGAGATA	TGGTGGACGA	2040
	AGGAATTTTA	TAACCTGTAC	AATTAAGGAT	TTTTATTATG	TGTTATTAT	TATTTCTTTT	2100
	TTGTGTTTGA	CTGCACAGGA	TCAAAATGCC	TGTTATCTCC	CTTTACTGGS	GACTTTTTTT	2160
25	TTTTTTTTTT	TTTTTTTTTA	TCAGACAGGG	TCTTGTCTTG	TGCCCCAGGC	TGGAGTGCAG	2220
	TGGTGGCATC	TGGGCTCACT	GCAACTTCAG	CTCTCTGGAT	TCAGGCAACA	CTCCTGCCTC	2280
	AGCCTCCAC	GTGGCTGGGA	TTACAGGTGC	CTGCCCCCAT	GGCTAATTTT	TTGTATTTTT	2340
	TGTAGAGATG	GGGTCTTACC	ATGTTGGCTG	GGCTGGTCTC	ACTCTCCCTGA	CCTCAAGCAA	2400
	TCTGCGCTGC	TCAGCCTCCC	AAAGTGCCTG	GATTACAGGC	GTGAGCCACC	GGCCCCAGCC	2460
30	TGAGCCTTTT	TTTTTTTCTA	ATGCATCCAA	GGTTAAGGGG	AAGACGCAAA	TAACAGGACT	2520
	ATTCTAAAG	GAACCTGTT	TGAACCTGTT	GAGATCAGTC	ATCAGTCTCA	GTATTCACCA	2580
	GGCACACCTT	AATTTTCATG	TAAAGAGATA	TATATATTTT	GTCTATTTTT	GTGCTTTTTG	2640
	GGGCTTATTT	TGTGCTTTTT	TACCTTATGT	AGAGATCTTA	TTACAAAGTG	ATTTCTTACA	2700
	TTAAAGAGAG	ACTGAATATA	ATTGTATAGT	TACTTAACTA	ATGAAGACAT	TTCAAGAACT	2760
35	TGGGATGATT	TTAATCTTGA	AGTAGTAGGT	GGTATAGTCA	TAAACCAAT	CATCCCTTTC	2820
	TTGATGTGAT	CTAATTTTTT	TGGCTTTAAG	GTGACATCTG	AGAGGTAAAG	CATTTCTTTT	2880
	TATATTGAAA	TCATAAACTA	TCACCGCGTG	CTTCTCTGAG	TTACTTTTAA	TTTTGCGCTG	2940
	TGTTATGGTT	TGGGCTTTTC	CTTCTGTTTG	GTTTTCAGAG	CCCCATGTCT	ATATAGTCTT	3000
	GAGTGCAGAT	AATCTCTATA	CTTGTAATAG	AAGATCAGTA	TTTCTGCCTA	GATCTGATAA	3060
40	AAAAATTTT	TGAATCTTGA	TATAAAATTT	CAAGAAATG	TGTTACAAAG	ATACTTAGTA	3120
	TAGCTCTTCA	GCATATACCT	GAGACTTGGG	ATGAATTTTA	AACCGAGATC	GATTTACTTT	3180
	GCAGATCAT	AGGCTTTTTA	TACTCTTGT	ATCAAAATGG	CTTATTTTTC	AGGCCTAAG	3240
	GATTTGTAAG	AGAAAGGCTT	TTCAACGAAAG	GATTCCTTTT	CTTCTCCCAT	ACTGTTCTTG	3300
	ATTTCTCTCT	TCCTTCAGGC	CTCAACAGGC	ACTGTATTTA	TTGCCAATGT	TCCAAATAT	3360
45	CAAAATTCAG	TGAATCTTAT	TGTGTTTCT	TTACTTATAT	AAAAAAGAT	AACCTTAAGG	3420
	ATGTGCAAGT	ACATTTCCAA	CTGCTAGCAC	AACCAGTATT	TTGTAAATTA	ACAAATCGCT	3480
	GTATGGTATG	GTCTTCTACA	CATTATATGC	TATAGATATC	TATCGATCAT	CTTTCTATTC	3540
	TGTTTCATGA	CTGAATAATG	TAAACCCAGT	GTTGGCAATT	GGTATCATCA	ATGATACTCA	3600
50	TTTTTTAATA	ACCAAGGCA	GGGGAATAAT	ATTTTACTTA	TTAATAAATA	TTTTATGATG	3660
	TGAAAAAATA	AAAAAATA	AAAAAATA				

A175 Protein sequence

Gene name: EBTs
 Unigene number: Hs.128899
 Protein Accession #: none found
 Signal sequence: 1-11
 Transmembrane domains: 402-424
 Cellular localization: not determined

60	1	11	21	31	41	51	
	MLSGFLMSPS	TQHRAYQTPG	GKFLPWEAS	GAHSTRGRGS	DRERESRPEA	AGLLNDRAAA	60
65	GEAREKGNRGE	PPAWIRAOQQ	PRPPFAGQAP	GTAAGGAQDP	RLRFGSRGR	VRLPVKPPEA	120
	SGRQPRGSPD	CIPRFPSASA	TEKAVEKGTG	PFASDGDGLG	APGFARRRR	LLGVAAEGSG	180
	PRGKRRTVSE	DEARGSPGPR	LLGURPALSG	DALGAPRVVP	CGAIAARPSP	HPGTPLRSCS	240
	CCWLRCWRRG	RQPSGEYCHG	NLDAQGVWRI	GFQCFERFDG	GDATICCGSC	ALRYCCSSAE	300
	ARLDQGGCDN	DRQGGAGEPG	RAEDQGPRL	GRASCLRGTD	GDGEGAPPFV	RAWQRCSPEG	360
70	SPRGRQLLRA	PGQLLPARR	RGFPSSFRGG	PSFLQKPALE	IYVEFLIVGS	VFVARILIGS	420
	IYVACCCCL	RPKQDPQQR	APGGRNLMET	IFMIPASATS	RGSSSRQSS	AASSSSSANS	480
	GARAPPTRSQ	TNCLPEGTN	NNVYVNMPTN	PSVINCQCAT	QIVPEQGQYL	HPFYVGYTVQ	540
	HDSVMTAVP	PFMDGLQPGY	RQIQSPFPRT	NSEQMTYPAV	TV		

A176 DNA SEQUENCE

Gene name: EBTs, Weakly similar to CGHUTL collagen alpha 1(III) chain precursor [H.sapiens]
 Unigene number: Hs.19322
 Probeset Accession #: AA088458
 Nucleic Acid Accession #: AA088458
 Coding sequence: 862-1995 (underlined sequences correspond to start and stop codons)

80	1	11	21	31	41	51	
	GCCCTTGGAC	ACTGACATGG	ACTGAAGGAG	TAGAATGGAG	CACGAGGACA	CTGACATGGA	60

5	CTGAAGAAAA	AGSAGCTGGA	GCAGGAGAAG	GAGGTGCTGC	TGCAGGGTTT	GGAGATGATG	120
	GGCGGGGGCC	GCAGCTGGTA	CCAGCAGCAG	CTGCACAGAG	TGCAGGAGCG	CCAGCGCCGC	180
	CTGGGGCCAGA	GCAGAGCCAG	CGCCGACTTT	GGGGCTGCAG	GGAGCCCCCG	CCCACTGGGG	240
	CGGCTACTGC	CCAGGTIACA	AGAGGTGGCC	CGGTGCCTGG	GGAGCTGCT	GGCTGCAGCC	300
	TGTGCCAGCC	GGGCGCTGCC	CCCGTCTCC	TCCGGGCCCC	CCTGCCCTGC	CCTGACGTCC	360
	ACCTCACCCC	CGGTCTGGCA	GCAGCAGACC	ATCCTCATGC	TGAAGGAGCA	GAACCGACTC	420
	CTCACCCAGG	AGGTGACCGA	GAAGAGTGAG	CGCATCACGC	AGCTGGAGCA	GGAGAGTCCG	480
	GCCTCATTA	AGCAGCTGTT	TGAGCCCTGC	GCCCTGAGCC	AGCAGGACGG	GGGACCTCTG	540
10	GATTCCACCT	TCATCTAGTC	CTTGTGGGCC	CGGTGGGCC	CCAGGGCCAG	CCTGGCACTC	600
	AGCCCTTCCA	GGGTGGGGCC	CCCATCGCAC	CCACCTCTCT	TGGCTGGAGA	CCCCCGGCAG	660
	GCCCAGGCAC	AGTCCCGGAG	TGGGGCGCTT	CTTGCOCGCC	TTGCCAGATG	GGCTCCCGAG	720
	GGCTGCCCCC	AGCTGGTCCC	CGACCCAGCC	GCTTGACTCC	GTTTTGGCTC	CTGGTTGYTG	780
	ACATGGGCTG	GGGGTCTCTG	TGAGTCCGCA	TAGTCCGCG	CTACTACTGG	CCGCTGTCTG	840
15	TGACAGTGG	GGTACCCCTC	CATGAGTLAG	CGTCCGCC	TTTCCAGCG	TGCCGCCCTG	900
	GGTCCCATCT	TACGGGAAAG	GCATGCCCCA	CGCCAGGCTG	CACTTCCAAC	AACGGGCAGC	960
	AGAGGGCGCG	GGGCGGCTCC	GACCGGGGTC	CAAGGGCAGC	TTCCCGCTCA	ACCAGGGCAC	1020
	CAGGACGAGG	TGGCTGTAGC	TCCGACCGAC	GGAGGTAGAT	GGAGGGGGTG	GGGACGGCCT	1080
	GTAAGCGGGG	GGTGCTTGCC	TGGCTGGGGA	GCCCCAGGGA	TAGCGGTCCG	ACTTCAGGTT	1140
20	CTGGCCAGG	CTGAGGTACC	CTGGCTGCAG	CGGATCGGCA	CGCCGGGTGG	GCGAGAGCTT	1200
	GGCTGCATG	TGCCCTCCAC	AGACCTGGG	GTGATGGCCT	TCCCCCTCTT	GGCCGGGACG	1260
	TTGCCCCACG	TTGAGTCCCA	CACACATCC	TGTGAGCCTG	GCTCCCCAGG	AGGGCCCCCA	1320
	GACAGCTCCC	AGCTGGTCCC	CGACCCAGCC	CTGTTTCCCC	CGACTCAGGA	TTTCCAAGGC	1380
	CTGGGGTCC	GGTCAACCCC	CTTGTCTCTC	ACGCCAGGCC	TGTCCCCAGG	TTTCAGCTGG	1440
25	GAGAGGCCAC	CTCCCTCAGC	CAAGGAAAC	GAGAACCCCC	AGGGTACAGG	AGGAGGCTGG	1500
	GGCAGGTCCC	CTTGGGTGTC	ACTCCTCAG	CCCCCTGCCA	GGCCCACTCC	CGTGGTGGCT	1560
	GGAGTACGCA	CTGGTGGGGG	GGCCCTGCTC	AGCCCAACCT	GGAGGGTCCC	AGTGTACCCA	1620
	GAACAGGGGG	CAAGCAACCA	GCATCGATGG	GTTCCTGAGC	CCAGGGCCCC	CGATGCGGGG	1680
	TCAGTGTGTG	TGGGGCGCAG	GGCCTCCGAT	GCGGGGTGAG	TGCGTGGGGG	GCGCAGGGCC	1740
30	CCGATGTGG	GGTCACTGCG	TGGGGGGCGC	AGGGCCCCCT	CGTGTCCAGG	GCACCTTGCT	1800
	ACACTGTCCC	ACAGGGCACC	TGTCTCAGAG	GAGGGGCCCT	GGCAGGCAGC	GTGGCACTC	1860
	CCCTTCGGAG	CCAGCTCCA	TGCTAACCTG	CCCAACAGCA	CCCCACAGAG	CCACATTCCC	1920
	TGCTGCACCT	GGTCTGCAGG	GGTGTCCAG	GACAGGCCCA	AGTCAGGCCA	GCATGCAGCT	1980
	GCCTCTCTAC	CTTGAAGATG	GGATGGGCT	TTCCAGGGGA	CATAAGGATG	TCAGGCTCTG	2040
35	ACCTCTCTGG	CTAGAAAGGG	CTAGGCTCCT	GAGGGCCTGT	GCCCCACAGC	CGCCACCCC	2100
	AGGTGGACTG	CAGCGCAGTG	GGTGGGCCAG	TGGCAGCCAG	GGAGAAAGCC	CCCGTCAGCA	2160
	GGCTGGGGTC	TGCCCAACAG	GGCCTCCCCA	CGTCTGCCTT	TGAGGGTGGC	TGCCATGCCC	2220
	TGGGGGATCC	CTTGGAGCCT	ACTGGACTGG	AAGCAGGAGA	CAGAACAGTG	TCTGTCCCGG	2280
	GGTGACTTCA	TCAGGAGACC	GCCCCACATG	AGCTGGACCC	CGCAGCTGAA	GCGGAAATGT	2340
40	GAGACAGCAC	TGTGATGACA	CCGGAACCTG	CTTTCAGCCT	TGGTGTCCG	TGCAGGTGA	2400
	AAAGAAATAG	GTCTCTCCAG	TTTACAGCTT	GAATCAGGC	TAGTGAGTGG	CCCTGGAGAC	2460
	CACGAGGGGA	GAATTTAAAG	GCCCCGGCTG	GCAGGGTCTA	GGTGGCTGGC	AGAGGCACAT	2520
	GCAGACCCCT	CTTGGAGCCT	GCCCCTAGGAC	GCTGGGCGGG	TCAGTCTCCG	TGCAGGATGT	2580
	GAGCAGCGTC	CTTGGGCTCT	ATCCGCGAGG	TGCCAGTAGC	GTGTGCAGGT	ACATACACGT	2640
45	GCTGTGCAC	TGTGATGACA	CCCGAAATG	TCTCAGGATG	TTGAATGTG	TCTTGGGGG	2700
	CAGAGGTGTC	CCGAGTTGAG	AATCTGCCCC	AGAGGAACAC	ACCCACACCA	GGCCTCAGGA	2760
	TTTTGTGTG	ATCAAGTTCC	AAGGAAAGG	AACATCTCAG	CCGGGCGTGG	TGGTTACGCG	2820
	CTGGAAATCC	AGCATTCTGAG	GCCAGGAGTT	CCAGAGCAGC	CTGGGCAACG	CAGTGAGAGA	2880
	CCCATCTCT	ACAAAGAAAA	AAAAAGAAAG	AAAGAAATG	AGAGATCCAG	GTTTAAAAAT	2940
50	TCATAAACAC	CACAAAGGAA	CAATACACTA	TGAGACCCAG	CAGAAAGCAAC	AGATTGACTC	3000
	TAGACCCAGA	TACTAGAAAT	ATCAGAGAGA	ATATAAGSTA	ACAGTGTTTT	ATATATCTAA	3060
	AGAAATAAAA	GAGATTCTCT	GAACATGAA	AAAAAA			

A177 DNA SEQUENCE

55	Gene name:	Cadherin 3, P-cadherin (placental)
	Unigene number:	Hs.2877
	Probeset Accession #:	X63629
	Nucleic Acid Accession #:	X63629
	Coding sequence:	54-2543 (start and stop codons are underlined)

60	1	11	21	31	41	51	
	GCAGAACACC	GGCCCGCCGT	CGCGGAGGCT	GCTTCACCCC	TCCTCTGCA	GCCATGGGGC	60
	TCCTCTGTCG	ACCTCTCGCG	TCTCTCTCTC	TTCTCCAGGT	TTGCTGGCTG	CAGTGCCTGG	120
65	CCTCCGAGCC	GTCCCGGGCG	GTCTTCAGGG	AGGCTGAAGT	GACCTTGGAG	GCGGGAGGGG	180
	CGGAGCAGGA	GCCCGGCCAG	GCCTGGGGA	AAGTATTTCAT	GGGCTCCCTT	GGGCAAGAGC	240
	CAGCTCTGTT	TAGCACTGAT	AATGATGACT	TCACTGTGGG	GAATGGCGAG	ACAGTCCAGG	300
	AAAGAAGGTC	ACTGAAGGAA	AGGAATCCAT	TGAAGATCTT	CCCATCCAAA	CGTATCTTAC	360
	GAAGACACAA	GAGAGATTGG	GTGGTTGCTC	CAATATCTGT	CCCTGAAAAT	GGCAGGGGTC	420
70	CCTTCCCCCA	GAGACTGAAT	CAGCTCAAGT	CTAATAAAGA	TAGAGACACC	AAGATTTTCT	480
	ACAGCATCAC	GGGGCCGGGG	GCAGACAGCC	CCCTGAGGG	TGTCTTCGCT	GTAGAGAAAG	540
	AGACAGGCTG	GTGTGTGTTG	AATAAGCCAC	TGGACCGGGA	GGAGTTTGCC	AAGTATGAGC	600
	TCTTTGGCCA	CGCTGTGTCA	GAGAAAGGTG	CCTCAGTGGG	GGACCCCATG	AACATCTCCA	660
	TCATCTGTAC	CGACCAAGAT	GACCAAGCC	CCAAGTTTAC	CCAGGACACC	TCCGAGGGGA	720
75	GTGTCTTAGA	CGAGTCTCTA	CCAGGTACTT	CTGTGATGCA	GGTGACAGCC	ACAGATGAGG	780
	ATGATGCCAT	CTACACCTAC	AATGGGGTGG	TTGCTTACTC	CATCCATAGC	CAAGAACCRA	840
	AGGACCCACA	CGACCTCATG	TTACACATTC	ACCGGAGCAC	AGGCACCATC	AGCGTCATCT	900
	CCAGTGGCCT	GGACCGGGAA	AAAGTCCCTG	AGTACACACT	GACCATCCAG	GCCACAGACA	960
	TGGATGGGGA	CGGCTCCACC	ACCAAGGCAG	TGGCAGTAGT	GGAGATCCCT	GATGCCAATG	1020
80	ACAATGTGCC	CACTGTGTAC	CCCGAGAGT	ACGAGGCCCA	TGTGCTGTAG	AATGCAAGTG	1080
	GCCATGAGGT	GCAGAGGCTG	ACGGTCTCTG	ATCTGGACGC	CCCCCACTCA	CCAGGCTGGC	1140
	GTGCCACCTA	CCTTATCATG	GGCGGTGACG	ACGGGGACCA	TTTTACCATC	ACCACCCACC	1200
	CTGAGAGCAA	CCAGGGGCATC	CTGACAAACA	GGAGGGGTTC	GGATTCTTGG	GCCAAAGACC	1260
	AGCACACCC	GTACGTTTGA	GTGACCAACG	AGGCCCTCTT	TGTGCTGAG	CTCCCAACCT	1320
	CCACAGCCAC	CATAGTGGTC	CACGTGGAGG	ATGTGAATGA	GGCACCTGTG	TTTGTCCACC	1380

5
10
15
20
25
30

```

CCTCCAAAGT CGTTGAGGTC CAGGAGGGCA TCCCCACTGG GGAGCCTGTG TGTGTCTACA 1440
CTGCAGAGA CCTGACCAAG GAGATCAAAA AGATCAGCTA CCGCATCCTG AGAGACCCAG 1500
CAGGGTGGCT AGCCATGGGAC CCAGACAGTG GGCAGGTGAC AGCTGTGGGC ACCCTCGACC 1560
GTGAGGATGA GCAGTTTGTG AGGACAACA TCTATGAAGT CATGGTCTTG GCCATGGACA 1620
ATGGAAGCCC TCCCACCACT GGCACGGGAA CCCTTCTGCT AACACTGATT GATGTCAACG 1680
ACCATGGCCC AGTCCCCTGAG CCCCCTCAGA TCACCATCTG CAACCAAAGC CCTGTGCGCC 1740
ACGTGCTGAA CATCACGGAC AAGGACCTGT CTCCCCACAC CTCCCCTTTC CAGGCCCGAC 1800
TCACAGATGA CTACAGATC TACTGGACGG CAGAGGTCAA CGAGGAAGGT GACACAGTGG 1860
TCTTGTCCCT GAAGAAGTTC CTGAAGCAGG ATACATATGA CGTGACCTT TCTCTGTCTG 1920
ACCATGGCAA CAAAGAGCAG CTGACGGTGA TCAGGGCCAC TGTGTGCGAC TGCCATGGCC 1980
ATGTCGAAGC CTGCCCCTGGA CCTGGGAAG GAGGTTTCAT CCTCCCTGTG CTGGGGGCTG 2040
TCCTGGCTCT GCTGTTCTCT CTGCTGGTGC TGCTTTTGT TGTGAGAAAG AAGCGGAAGA 2100
TCAAGGAGCC CTTCTACTC CCAGAAGATG ACACCCGTGA CAACGCTTTC TACTATGGCG 2160
AAGAGGGGGG TGGCGAAGAG GACCAAGACT ATGACATCAC CCAGTCCAC CGAGGTCTGG 2220
AGGCCAGGCC GGAGGTGCTT CTCCGCAATG ACGTGGCACC AACCATCATC CCGACACCCA 2280
TGTACCGTCC TAGGCCAGCC AACCAGATG AATCGGCAA CTTTATAAT GAGAACCTGA 2340
AGGCGGCTAA CACAGACCCC ACAGCCCGCG CCTACGACAC CCTCTTGGTG TTCGACTATG 2400
AGGGCAGCGG CTCGCGAGCG GCGTCCCTGA GCTCCCTCAC CTCTCCGCC TCCGACCAAG 2460
ACCAAGATTA CGATTACTG AACGAGTGGG GCAGCGCTT CAAGAAGCTG GCAGACATGT 2520
ACGGTGGCGG GGAGGAGCAG TAGCGGCGCT GCCTGCAGGG CTGGGGACCA AACGTCAGGC 2580
CACAGAGCAT CTCCAGGGG TCTCAGTTCC CCTTCAGCT GAGGACTTCG GAGCTTGTCA 2640
GGAAATGGCC GCTGCAACTT GCGGAGAGA GGCATAGAT CTGACGTTAG AGTGGTTGCT 2700
TCTTAGCTCT TCCAGGATG AGGAATGTGG GCAGTTGAC TTGAGCACTG AAAACCTCTC 2760
CACCTGGGCC AGGCTTCCCT CAGAGGCCAA GTTTCCAGAA GCCTCTTACC TGCCTGAAA 2820
TGCTCAACCC TGTGTCTCTG GCCTGGGCGT GCTGTGACTG ACCTACAGTG GACTTCTCT 2880
CTGGAATGGA ACCTTCTTAG GCTCCTGCTG GCAACTTAAT TTTTCTTTT AATGCTATCT 2940
TCAAAACGTT AGAGAAAGT CTCAAAAGT GCAGCCCGAG CAGCTGGGCG CCACTGGCGC 3000
TCTGTGATTT CTGCTTTCCA GACCCCAATG CCTCCCATTC GGTATGATCT CTGCGTTTTT 3060
ATACTGATG TGCTAGGTT GCCCTTATT TTTTATTTT CCGTGTGCGT TGCTATAGAT 3120
GAAGGGTGAG GACAATCGTG TATATGTACT AGAATTTTT TATTAAGAA A

```

A178 Protein sequence:

35
40

```

Gene name: Cadherin 3, P-cadherin (placental)
Unigene number: Es.2877
Probeset Accession #: X63629
Protein Accession #: CAA45177
Signal sequence: 1-24
Transmembrane domain: 659-675
Cellular localization: plasma membrane

```

45
50
55
60

```

1 11 21 31 41 51
MGLPRGPLAS LLLLVQCNLQ CMASEPCRAV FREAEVTLFA GGABQEPGQA LGKVFMGCPG 60
QEPALPSTON DDFTVENGST VQERRSLKER NPLKIPPSKR ILRRHKRWV VAPISVPENG 120
KGFPPQRLNQ LKSNKDRDTK IFTSITGPGA DFPPEGVFAV EKETGWLILN KPLDREHIAK 180
YELFGHAVSE NGASVEDEMN ISIIVTDQND HKFKPTQDTF RGSVLGVLV GTSVMQVTAT 240
DEDDAITYTN GVVAYSIESQ EPKDPHDLMF TTHRTGTIS VISSGLDREK VPEYTLTIOA 300
TMDGDSST TAVAVVSLD ANDMAEMFDP QKYRAHVPEV AVGERVQRLT VTDLDAFNSP 360
AWRATYLING GDDGDHFTIT THPESNQIIL TTRKGLDFEA KNQHTLYVEV TNEAPFVLKL 420
PTSTATIVVH VEDVNEAPVF VPPSEKVEVQ EGIPTGEPVC VYTAEDPDKK NQKISYRILR 480
DPAGWLAMD PDSGQTVAGT LDREDEQFVR NNIYEVMLA MDNGSPPTTG TGTLLTLTLD 540
VNDHGPVPEP RQITICNQSP VRHVLNITDK DLSPTSTPFO AQLTDDSDIY WTAEVNEEGD 600
TVVLSLKKPL KQDTYDVHLS LSHKGNKEQL TVIRATVCD C HGHVETCPGP WKGGFILFVL 660
GAVLALLFL LVLVLLVRK RKIKEPLLLP EDDTRDNVYF YGREGGGEED QDYDITQLER 720
GLBARPEVVL RNDVAPTII P TMYRFRFPAN PDEIGNFIIE NLKAANTDPT APPYDTILNF 780
DYSGSGSDAA SLSSITSSAS DQDQDYDILN EWGSRFKKLA DMYGGGEED

```

A179 DNA SEQUENCE

65

```

Gene name: endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
Unigene number: Es.258583
Probeset Accession #: NM_012152
Nucleic Acid Accession #: NM_012152
Coding sequence: 43-1104 (underlined sequences correspond to start and stop codons)

```

70
75
80

```

1 11 21 31 41 51
CTTCTTTAAA TTCTTTCTA GGAATGTCAC TTCTTCTCCA CAATGAATGA GTGTCACTAT 60
GACAAGCACA TGGACTTTTT TTATAATAGG AGCAACACTG ATACTGTCCA TGACTGGACA 120
GGAACAAGGC TTGTGATTGT TTGTGTGTTT GGGACGTTTT TCTGCTGTTT TATTTTTTTT 180
TCTAATCTCT TGTGATATGC GGCAGTGATC AAAAACAGAA AATTTCAATT CCCCTTCTAC 240
TACCIGTGTG CTAATTAGC TGCTGCCGAT TTCTTCTGCT GAATGCCCCA TGTATTCCAG 300
ATGTTTAAAC CAGGCCAGT TTCAAAACT TTGACTGTCA ACCGCTGGTT TCTCCGTCTG 360
GGGCTTCTGG ACAGTAGCTT GACTGCTTCC CTCACCAACT TGCTGCTTAT GCGCGTGGAG 420
AGGCACATGT CAATCATGAG GATGCGGGTC CATAGCAACC TGACCAAAA GAGGGTGACA 480
CTGCTCATTT TGCTTGTCTG GGCATCGGCC ATTTTATAGG GGGCGGTCCC CACACTGGGC 540
TGAATTTGOC TCTGCAACAT CTCTGCCCTG TCTTCCCTGG CCCCATTATA CAGCAGAGAT 600
TACCTGTGTT TCTGGACAGT GTCCAACTTC ATGGCCCTTC TCATCATGTT TGTGGGTGAC 660
CTGCGATTTT ACCTGTACGT CAAGAGGAAA ACCAAGCTCT TGTCTCCGCA TACAAGTGGG 720
TCCATCAGCC GCGCGAGGAC ACCCAAGAG CTAATGAAGA CGGTGATGAC TGTCTTAGGG 780

```

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

GGGTTTGTGG TATGCTGGAC CCCGGCCCTG GTGGTTCTGC TCCTCGACGG CCTGAAGTGC 840
 AGGCAAGTGTG GCGTGCAGCA TGTGAAAAGG TGGTTCTCTG TGCTGGCGCT GCTCAACTCC 900
 GTCTGAACCC CCATCATCTA CTCTACCAAG GACGAGGACA TGATATGGAC CATGAAGAAG 960
 ATGATCTGCT GCTTCTCTCA GGAGAACCAC GAGAGGCGTC CCTCTCGCAT CCCCTCCACA 1020
 GTCTCTAGCA GGAGTGACAC AGGCAGCCAG TACATAGAGG ATAGTATTAG CCAAGGTGCA 1080
 GTCTGCAATA AAAGCACTTC CTAAACTCTG GATGCCCTCTC GGCCACCCCA GGTGATGACT 1140
 GTCCTTAGG

A180 Protein sequence:

Gene name: endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 7
 Unigene number: Hs.258583
 Probeset Accession #: NM_012152
 Protein Accession #: NP_036284
 Signal sequence: none found
 Transmembrane domains: 31-53, 66-88, 150-172, 190-211, 239-261, 277-295
 Cellular Localization: plasma membrane

1 11 21 31 41 51
 MNECHYDKHM DFFYNRSNTD TVDDWTGTXL VIVLCVGTFF CLPIFFSNLS VIAAVIKNRK 60
 FHPFFYLLA NLAAADFFAG IAYVFLMFT GPVSKILTIN RWFRLQGLD SSLTASLTNL 120
 LVIAVERHMS IMRMVRHNSL TGRVTLILL LVMAIALEMG AVPTLGWNCL CNISACSSLA 180
 PIYSRSLVLF WTSYNTLMAPL IMVVVYLRIY VYVKRKNVL SPRTSGSISR RRTPMKLMKT 240
 VMTVLGAPVV CMTPGLVLL LDGLNCRQCG VQHVWRWFL LALLNSVVP IYYSYKQEDM 300
 YGTMMKMTCC PSQENKERRP SRIPSTVLSR SDTGSQYIED SISQGAVCNK STS

A181 DNA SEQUENCE

Gene name: ESTs
 Unigene number: Hs.162859
 Probeset Accession #: AA569531
 Nucleic Acid Accession #: AA569531
 Coding sequence: 1-504 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 ATGACCTACA GTTACTCAT TTTCAAGGCT GAGTGTATCG TTAATCATCT TAATTATGTT 60
 CATTCGTGAAG CCACACAGGAG AACCAAGACC AAAACTTTAT TGCTCTGCT TCAATTCTCT 120
 GATGAACCT CTGACATAAG CACACATCTT CCTGTGTTAT CTCTCTCAA GGAGTGTGGA 180
 GTGCTTCATC TGCATATCCA CGGGAAGAAG GAAGACATGA GAATCACCCA ACAGTCTTCC 240
 CAGCTATAAC TGTTGGGACAT GGGTGGTTTT ACAATATTTA AGAACCTGTG GATGAGCCTC 300
 ATACCCAGAG GGAACAAACG CTCCCCAAAA AGAGTTACAG AAACCATCCT GAGAGATTTT 360
 AAGCAGAAGC AAGTTCAAAA GATCCCAAGAG GAGAGACGAA GAGAGTCTGC AGGACCAAAC 420
 CTCTCTCAT TCTGTTTGT GGGGAATGCT GGAAGAGGAG ACAGGCCCCA GATTTGGGCA 480
 GGAAGTAAAC AGTTTTCAGG CTGAGGCCAA TCTGAGCAG AACATTCCAA TATTTCTTCA 540
 GCTACGTTGT CCCAGCACCT CACTGGTTAA CCTTTATGT CCACCAATTG TGGATTTCAC 600
 AGCTACTTGT CAATGGTGAA TATTGATCAT CATCATATC TACTGAGCTG CTACCATATC 660
 CCAGCTACTC CTGTCATGTT GTTCATTATT TTCTCAACAC TCAGCATATT TGCAATATGT 720
 TATGTAATAT CACAGACAG GAAACTGAAC GCAGAAATGT TTATTTCTT GCCAAACATC 780
 ACATGAGGAT GAACAATGAA ACGATTGTA AACCCAGGAT GTCTGATTCC AACATCTCTG 840
 GGTCTTTTTT CACTCTGATA TGCTGCAATT AAAAGCCAT TTCTAAGACT GT

A182 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.162859
 Probeset Accession #: AA569531
 Protein Accession #: none found
 Signal sequence: 1-46
 Transmembrane domains: none found
 Cellular Localization: not determined

1 11 21 31 41 51
 MTYSYSFFRP ELIVNHLNYV HSEANRRYKT KTLNLLSFL DETSGLSLHL PCLSLSKKCG 60
 VLHLDIHGKK EDMRTQQSS QLYLWDMGCF TIFENLWMSL IPRGNKRSFK RVTETILRDF 120
 KQKQSSKIQE ERRRESAGFN LSSFWFVQNA GRGDRPQIWA GSKQFSG

A183 DNA SEQUENCE

Gene name: ESTs
 Unigene number: Hs.179809
 Probeset Accession #: N95796
 Nucleic Acid Accession #: XM_050197
 Coding sequence: 310-1971 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 TCACAGTGC CAAGGGGCTG GCTCAGCGGA ACCAGCCTGC ACCGCTGCG TCGGGGTGAC 60
 AGCGCGCGCG CTCGGCCAGG ATCTGAGTGA TGAGACGTGT CCCACTGAG GTGCCCCACA 120
 GCAGCAGGTG TTGAGCATGG CCTGAGAAGC TGGACCGGCA CCAAGGGCTG GCAGAAATG 180
 GCGGCTGCG TGAATCTTAG CGAGTTGGCG GCAGCAAGGA GGAGAGGCGG CAGCTTCTCG 240
 AGCAGAGCCG AGACGAAGCA GTTCTGGAAT GCCTGAACGG CCCCTGAGC CCTACCGGCC 300

5 TGGCCCACTA TGGTCCAGAG GCTGTGGGTG AGCCGCTGCTG TCGGACACCG GAAAGCCAG 360
 CTCTTGCTGG TCAACCTGCT AACCTTTGGC CTGGAGGTGT GTTGGCCGCG AGGCATCACC 420
 TATGTGCCGC CTCTGCTGCT GGAAGTGGGG GTAGAGGAGA AGTTTCATGAC CATGGTGCTG 480
 GGCAITGGTC CAGTGTGGGG CTGGTCTGT GTCCGCTCC TAGGCTCAGC CAGTGACAC 540
 TGGGTGGGAC GCTATGGGCG CGCGCGGCCC TTCTATCTGG CACTGTCTTT GGGCATCTCT 600
 CTGAGGCTCT TTCTCATCCC AAGGGCCGCG TGGCTAGCAG GGTCTGCTGT CCGGATCCC 660
 AGGCCCTTGG AGCTGGCCTG GCTCATCTCT GCGTGGGGC TGCTGGACTT CTGTGGCCAG 720
 GTGTGCTTCA CTCCACTGGA GGCCTGCTCT TCTGACTCTT TCGGGGACCC GGACCACTGT 780
 10 CGCCAGGCTT ACTCTGTCTA TGCCTTCATG ATCAGTCTTG GGGGCTGCTT GGGCTACCTC 840
 CTGCTGCTCA TTGACTGGGA CACCAAGTGC CTGGCCCCCT ACCTGGGCAC CCAGGAGGAG 900
 TGCTCTTTT GCTGCTCAC CTTCTCTTCT CTCACCTGCG TAGCAGCCAC ACTGCTGGTG 960
 GCTGAGGAGG CAGCGCTGGG CCCCACCGAG CCAGCAGAAG GGTGTGCTGC CCCCTCTCTT 1020
 TCGCCCCACT GCTGTCTATG CCGGGCCGCG TTGGCTTTCC GGAACCTGGG GCGCTGCTT 1080
 15 CCGCGCTGCG ACCAGCTGTG CTGCGGCTAG CCGCGCACCC TGGCGCGGCT CTTGCTGGCT 1140
 GAGCTGTGCA GCTGATGGC ACTCATGACC TTACAGCTGT TTACACGGA TTCTGTGGG 1200
 GAGGGCTGT ACCAGGCGT GCCCAGAGCT GAGCCGCGCA CCGAGGCGCG GAGACACTAT 1260
 GATGAAGGCG TTCCGATGGG CAGCCTGGGG CTGTTCTGCG AGTGGCCAT .CTCCCTGGTC 1320
 TTCTCTCTGG TCACTGACCG GCTGGTGCAG OGATTCGGCA CTCGAGCAGT CTATTTGGCC 1380
 20 AGTGTGGCAG CTTTCCCTGT GCTGCGCGT GCCACATGCC TGTCCACAG TGTGGCGGTG 1440
 GTGACAGCTT CAGCGCGCTT CACCGGGTTC ACCTTCTCAG CCTGCGAGAT CCTGCCCTAC 1500
 ACCTGCTGT CCGCTCTACA CCGGAGAGAG CAGGTGTCTC TGCCCAATA CCGAGGCGAC 1560
 ACTGGAGGTG CTAGCAGTGA GGACAGCCTG ATGACCACTG TCCTGCCAGG CCCTAAGCCT 1620
 GGGAGCTCCTT TCCCTAATGG ACACGTGGGT GCTGGAGGCA GTGGCTGCT CCCACCTCCA 1680
 25 CCGCGCTGCT AGGGGCGCTC TGCTGTGAT GTCTCGTAC GTGTGGTGTG GGTGAGGCC 1740
 ACCGAGGCCA GGGTGGTTC GGGCGGGGCG ATCTGCTGCG ACCTGCGCAT CCTGGATAGT 1800
 GCTCTCTGCT TGTCCAGGT GSCCCATCC CTGTTTATGG GCTCCATTGT CCAGCTCAGC 1860
 CAGTCTGTCA CTGCTATAT GGTGTCTGCC GCAGGCTGCG GTCTGGTCTC CATTTACTTT 1920
 GCTACACAGG TAGTATTTGA CAGAGCGGAC TTGGCCAAAT ACTCAGCGTA GAAACTTCC 1980
 30 AGCAGTGG GGTGGAGGCG CTGCTCACT GGGTCCAGC TCCCGCTCC TGTAGCTCC 2040
 ATGGGGCTG CCGGCTGGCC GCGAGTTCT GTTGTGCCA AAGTAATGTG GCTCTCTGCT 2100
 GCAACCTGT GCTGCTGAGG TGGTGTGCTG CACAGCTGGG GGTGGGGCG TCCCTCTCT 2160
 CTCTCCCGAG TCTTAGGGC TGCTGTGCTG GAGGCTTCC AAGGGGCTT CAGTCTGGAC 2220
 TTATACAGGG AGGCAGAGG GGTCTCATGC ACTGGATGCG GGGGACTCT CAGGTGGATT 2280
 35 ACCCAGGCTC AGGCTTAACA GCTAGCCTCC TAGTTGAGAC ACACCTAGAG AAGGGTTTT 2340
 GGGAGCTGAA TAACTCAGT CACTGGTGT OCCATCTCTA AGCCCTTAA CCTGCGAGCT 2400
 CGTTAATGT AGCTCTTGA TGGAGTTTC TAGGATGAAA CACTCTCTCA TGGGATTGA 2460
 ACATATGAAA GTTATTGTG GGGGAGAGT CCGAGGGGCG AACACACAG AACCGAGTCC 2520
 40 CCTCAGCCCC ACAGGCACTG GTCTTTTGTG CTNGANTCCA CCCCCCTCT CTTTACCTT 2580
 TT

A184 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.179809
 45 Probeset Accession #: N95796
 Protein Accession #: XP_050197
 Signal sequence: none
 Transmembrane domains: 21-43, 53-75, 90-112, 125-147, 160-182, 199-221, 323-345, 350-372, 379-401
 50 Cellular Localization: plasma membrane

55 1 11 21 31 41 51
 MVQRLWVSL LRHRKAQLL VMLTFQLEV CLAAGITYVP PLLLEVGVVE KFTMTVLGIG 60
 PVLGLVCPVL LGSASDHWRG RYGRRRPFIN ALSGLILLSL FLIPRAGWLA GLLCFDPRFL 120
 ELALLILGVG LDFQCVCF TPLBALLSDL FRDPDCEQRA YSVYAFMISL GGLGLYLLPA 180
 IDWDTSLALP YLGTQRECLF GLLTLIFLTC VAATILVAER AALQTEPAE GLSAPSLSPH 240
 CCPCRARLAF RNLGALLPRL HQLCCRMFRT LRRLFVAREL SNMALMTFTL FYTFDVGEG 300
 60 YQGVPRAPBP TRARRHYDEB VRMGSLGLFL QCAISLVFSL VMDRLVQRFQ TRAVYLAESA 360
 AFPVAAATC LSHSVAVVIA SAALTGFTFS ALQILPYTLA SLVHREKQVF LKPYRGDTGG 420
 ASSEDSLMTS FLGPKPGAP PFNGHVAGG SGLLPFEAL CGASACDVSV RVVVGEPTER 480
 RVVPERGICL DLAILDSAPL LSQVAPSLEM GSIVQLSQSV TAYMVSAGL GLVAITYFATQ 540
 VVFKSLLAK YSA

A185 DNA SEQUENCE

Gene name: ESTs
 Unigene number: Hs.11260
 Probeset Accession #: R73640
 70 Nucleic Acid Accession #: AK002126
 Coding sequence: 1-1593 (underlined sequences correspond to start and stop codons)

75 1 11 21 31 41 51
 ATGGTTGCCC GGGGGCTGCT TGGTGGATT TCCCGGGTGG TGGTTTGTCT GGTGCTCTCT 60
 TGTGTGCTA TCTCTGTCTT GTACATGTG GCTGTCACCC CAAAGGTGA CAGGAGCAG 120
 CTGGCACTGC CCAGGGCCAA CAGCCCCACG GGAAGGAGG GGTACCAGC CGTCTCTCAG 180
 GAGTGGGAGG AGCAGCACCG CAACTACGTG AGCAGGCTGA AGCGGCAGAT CGCAGAGCTC 240
 80 AAGGAGGAGC TSCAGGAGAG GAGTGAGCAG CTCAGGAATG GGCAGTACCA AGCCAGGAT 300
 GCTGCTGCCC TGGGTCTGGA CAGGAGCCCC CCAGAGAAA CCCAGGCGGA CCTCTGCGCC 360
 TCTCTGCACT CGCAGGTGGA CAGGCAGAG GTGAATGCTG GGTCTCAGCT GGCACAGAG 420
 TATGCAGCAG TGCTTTTCTA TAGCTTTACT CTACAGAAGG TGTACCAGCT GGAGACTGGC 480
 CTTACCCGCC ACCCCGAGGA GAAGCCTGTG AGGAAGGACA AGCGGGATGA GTTGGTGGAA 540
 GCCATTGAAT CAGCCTTGGG GACCTGAGAC AATCTGCAG AGACAGGCC CAATCACCGT 600

5
10
15

```

CCTTACACGG CCTCTGATTT CATAGAAGGG ATCTACCGAA CAGAAAGGGA CAAAGGGACA 660
TTGTATGAGC TCACCTTCAT AGGGGACCAC AAACACGAAT TCACACGGCT CATCTTATTT 720
CGACCATTCG GCCCATCAT GAAAGTGAAA AATGAAAAGC TCACACATGGC CAACACGCTT 780
ATCAATGTTA TCGTGCCCTCT AGCAAAAAGG GTGGACAAAGT TCCGGCAGTT CATGCAGAAT 840
TTCAGGGAGA TGTGCATTGA GCAGGATGGG AGAGTCCATC TCACTGTGTG TTACTTTGGG 900
AAAGNAGAAA TAAATGAAT CAAAGGAATA CTGAAAACA CTTCCAAAGC TGCCAACTTC 960
AGGAACTTTA CCTTCATCCA GCTGAATGGA GAATTTTCTC GGGGAAAGGG ACTTGATGTT 1020
GGAGCCGCTT TCTGGAAGGG AAGCAACGTC CTCTCTTTT TCTGTGATGT GGACATCTAC 1080
TTCACATCTG AATTCCCTCA TACGTGTAGG CTGAATACAC AGCCAGGGAA GAAGGTATTT 1140
TATCCAGTTC TTTTCAGTCA GTACAATCTT GGCATAATAT ACGGCCACCA TGATGCAGTC 1200
CCTCCCTTGG AACAGCAGCT GGTCAATAAG AAGGAACTG GATTTTGGAG AGACTTTGGA 1260
TTTGGGATGA CGTGTCACTA TCGGTCAGAC TTCATCAATA TAGGTGGGTT TGATCTGGAC 1320
ATCAAAAGCT GGGCGGAGGA GGATGTGCAC CTTTATCGCA AGTATCTCCA CAGCAACCTC 1380
ATAGTGGTAC GGACGCTGT GCGAGGACTC TCCACCTCT GGCATGAGAA GCGCTGCATG 1440
GACGAGCTGA CCCCAGAGCA GTACAAGATG TGCATGCAGT CCAGGCCAT GAACGAGGCA 1500
TCCCACGGCC AGCTGGGCAT GCTGGTGTTC AGGCACGAGA TAGAGGCTCA CCTTGGCAAA 1560
CAGAAACAGA AGACAAGTAG CAAAAAACA TGA

```

20
25

A186 Protein sequence:
Gene name: ESTs
Unigene number: Hs.11260
Probeset Accession #: R73640
Protein Accession #: NP_060841
Signal sequence: 1-76
Transmembrane domains: none found
Cellular Localization: not determined

30
35
40

```

1 11 21 31 41 51
MVRRLGLANI SRVVVLLVLL CCAISVLVYL ACTPKGDEEQ LALFRANSPT GKEGYQAVLQ 60
EMEEQHRNVV SSLKRQIAQL KEELQERSEQ LRNGYQASD AAGLGDRSP PEKTQADLLA 120
FLHSQVDKAE VNAGVKLATE YAAVPFDSPT LQKVYQLETS LTRHPEEKV RKDKRDELVE 180
AIESALETLN WFAENSPNHR PYTASDFLEG IYRTERDKGT LYELTFKGDH KHEFKRLILF 240
RPFPGIMVK NEKLNMANTL INVIVFLAKR VDKERQFMQN FREMCIBQD RVHLTVVYFG 300
KEBINEVKSI LENTSKAAMF RNFTFIQNG EFSRQKGLDV GARFWKGSNV LLFFCDVDYI 360
FTSEPLNTGR LNTQPKKVF YPVLFSQYNP GIYGHEDAV PPLEQQLVIK KETGFWRDFG 420
FGMTQYRLSD FINIGFOLD IKWGGEDVH LYRKYLSNLS IVVRTPVRL FHLWHEKRCM 480
DELTPEQYKM CQSKAMNEA SHGQLGLVLF RHEIEHLRK QKQKTSKKM

```

45

A187 DNA SEQUENCE
Gene name: ATPase, Ca++ transporting, type 2C, member 1
Unigene number: Hs.106778
Probeset Accession #: N51919
Nucleic Acid Accession #: AF189723
Coding sequence: 1-2712 (underlined sequences correspond to start and stop codons)

50
55
60
65
70
75
80

```

1 11 21 31 41 51
ATGATTCCTG TATTGACATC AAAAAAAGCA AGTGAATTAC CAGTCAGTGA AGTTGCAGAC 60
ATTCCTCAAG CTGATCTTCA GAATGGTCTA AACAAATGTG AAGTTAGTCA TAGGCGAGCC 120
TTTCATGGCT GGAATGAGTT TGATATTAGT GAAGATGAGC CACTGTGGAA GAAGTATATT 180
TTCAGTTTGA AAAATCCCTT TATTATGCTG CTCTGGCTT CTGCAGTCAT CAGTGTTTTA 240
ATGCATCAGT TTGATGATGC CGTCAGTATC ACTGTGGCAA TACTTATCGT TGTACAGTT 300
GCTTTTGTTC AGGAATATCG TTCAAGAAAA TCCTTTGAGG AATTGAGTAA ACTTGTGCCA 360
CCAGAATGCC ATTTGTTGCG TGAAGGAAAA TTGGAGCATA CACTTGGCCG AGACTTGGTT 420
CCAGGTGAZA CAGTTTCCCT TTCTGTGGGG GATAGAGTTC CTGCTGACTT ACGCTTGTTT 480
GAGGCTGTGG ATCTTTCCAT TGATGAGTCC AGCTTGACAG GTGAGACAA CCGTTGTCTT 540
AAGGTGACAG CTCCCTCAGC AGCTGCAACT AATGAGATC TTGCATCAGG AAGTAACATT 600
GCTTTTATGG GAACACTGGT CAGATGTGGC AAAGCAAGG GTGTTGTCTT TGGAAACGGA 660
GAAAATTCCT AATTGGGGGA GGTTTTAAAA ATGATGCAAG CAGAAGAGGC ACCAAAAACC 720
CTCTGCAGA AGAGCATGGA CCTCTTAGGA AAACAACCTT CCTTTTACTC CTTTGGTATA 780
ATAGGAATCA TCATGTTGGT TGGCTGGTTA CTGGGAAAA ATATCTCGGA AATGTTTACT 840
ATTAGGTGTA GTTTGGCTGT AGCAGCAATT CCTGAAGGTC TCCCATTTGT GGTCAAGTGT 900
ACGCTAGCTC TTGGTGTAT GAGAAATGGT AAGAAAAGGG CCATTGTGAA AAGCTGCCT 960
ATTGTTGAAA CTCTGGGCTG CTGTAATGTG ATTTGTTGAG ATAAACTGG AACACTGAG 1020
AAGAATGAAA TGACTGTTAC TCACATATTT ACTTCAGATG GTCTGCATGC TGAGGTTACT 1080
GGAGTTGGCT ATAATCAATT TGGGGAAGTG ATTGTTGATG GTGATGTTGT TCATGGATT 1140
TATAACCCAG CTGTTAGCAG AATTGTTGAG GCGGGCTGTG TGTGCAATGA TGCTGTAATT 1200
AGAAACAATA CTCTAATGGG GAAGCCACCA GAAGGGGCCCT TAATTGCTCT TGCAATGAAG 1260
ATGGGTCTTG ATGGACTTCA ACAAGACTAC ATCAGAAAAG CTGAATACCC TTTTAGCTCT 1320
GAGCAAAAGT GGATGGCTGT TAAGTGTGTA CACCGAACAC AGCAGGACAG ACCAGAGATT 1380
TGTTTATGTA AAGGTGCTTA CGAACAGTA ATTAAGTACT GTACTACATA CCAGAGCAAA 1440
GGGCAGACCT TGACACTTAC TCAGCAGCAG AGAGATGTGT ACCAACAGA GAAGGCACGC 1500
ATGGGCTCAG CCGGACTCAG AGTTCTTCTT TGGCTTCTG GTCCCTGAAT GGCAGACCTG 1560
ACATTTCCTG GCTTGGTGGG AATCATTGAT CCACCTAGAA CTGGTGAGAA AGAAGCTGTT 1620
ACACACTCA TTGCTTCAG AGTATCAATA AAAATGATTA CTGGAGATT CAGAGGACT 1680
CGAGTTGCAA TCGCAGTGT TCTGGGATTG TATTCCAAAA CTTCAGCTC AGTCTCAGGA 1740
GAGAAATAG ATGCAATGGA TGTTCAGCAG CTTTCACAAA TAGTACCAA GGTTCAGTA 1800
TTTCAACAG CTAGCCCAAG GCACAAGATG AAAATTATTA AGTCGCTACA GAAGAACGTT 1860
TCAGTTGTAG CCATGACAGG AGATGGAGTA AATGATGAG TTGCTCTGAA GGCTGCAGAC 1920
ATTGGAGTTG CGATGGGCA GACTGGTACA GATGTTTGCA AAGAGGCAGC AGCATGTATC 1980
CTAGTGGATG ATGATTTTCA AACATAATG TCTCAATCG AAGAGGGTAA AGGGATTAT 2040

```

5
 10
 AATAACATTA AAAATTTCGT TAGATTCCAG CTGAGCACGA GTATAGCAGC ATTAACCTTA 2100
 ATCTCATTTG CIACATTAAT GAACCTTCCT AATCCTCTCA ATGCCATGCA GATTTTGTGG 2160
 ATCAATATTA TTATGGATGG ACCCCCAGCT CAGAGCCTTG GAGTAGAACC AGTGGATAAA 2220
 GATGTCAATC GTAAACCTCC TCGCAACTCG AAAGACAGCA TTTTGACTAA AAACCTGATA 2280
 CTTAAATAC TTGTTTCATC AATAATCAT GTTTGTGGGA CTTTGTTTGT CTCTGGCGGT 2340
 GAGCTACGAG ACATATGAT TACACCTCGA GACACAACAA TGACCTTCAC ATGCTTTGTG 2400
 TTTTITGACA TGTTCATGTC ACTAAGTTCC AGATCCCGA CCAAGTCTGT GTTTGAGATT 2460
 GGACTCTGCA GTAATAGAAT GTTTTGTCTAT GCAGTTCTTG GATCCATCAT GGGACAATTA 2520
 CTAGTATTTT ACTTCTCTCC GCTTCAGAAAG GTTTTTCAGA CTGAGAGCCT AAGCATACTG 2580
 GATCTGTGTG TTTCTTTGGG TCTCACCTCA TCAGTGTGCA TAGTGGCAGA AATTATAAAG 2640
 AAGGTTGAAA GGAGCAGGGA AAAGATCCAG AAGCATGTGA GTTCGACATC ATCATCTTTT 2700
 CTTGAAGTAT GA

15 **A188 Protein sequence:**
 Gene name: ATPase, Ca++ transporting, type 2C, member 1
 Unigene number: Hs.106778
 Probeset Accession #: N51919
 Protein Accession #: AAF27813
 20 Signal sequence: none found
 Transmembrane domains: 60-82, 86-108, 252-274, 282-304, 757-779, 827-849, 856-878
 Pfam domains: Cation ATPase N [6-83], E1-E2 ATPase [89-324], Hydrolase [328-649], Cation ATPase C [744-889]
 Cellular Localization: not determined

25
 30
 35
 40
 1 11 21 31 41 51
 MIPVLTSKKA SELPVSEVAS ILQADLQNGL NKCEVSHRRA FHGWNEFDIS EDEPLWKYKI 60
 SQFKNPLIML ILLASAVISVL MHQFDDAVSI TVAILIVTVV AFVQEYRSEK SLBELSKLVP 120
 PECHCVRECK LEHTLARDLV PGDTVCLSVG DRVPADLRLE EAVDLSDIES SITGETTPCS 180
 EVTAPQPAAT NGDILASRNI AFMGTLVRCG KARGVVGITG ENSEEGVEFK MMQAEAPKT 240
 PLQKEMDLLG KOLSPYSFGI IGIIMLVGWL LGKDILEMPT ISVSLAVAAI PEGPLIVTVV 300
 TLALGVMRMV KKRATVKKLP IVETLGCCNV ICSDRTGTLT KNEMTVTHIF TSDGLHAEVT 360
 GVGYNQFGEV IVDGDVVGHF YNPVAVSRIV AGCVQDAVI RNNTLMGKPT EGALIALAME 420
 MGLDGLQQDY IRKARYPFSS EQKMMAVKCV HETQQDRPEI CFMKGAYBQV IKYCTTFQSK 480
 35 GQTLTLTQQQ RDVYQOEKAR MGSAGLRVLA LASGPGLQL TELGLVGIID PPRTGVEKAV 540
 TLLIASGVSI KMITGDSQET AVAIASRLGL YSKTSQSVSG ZEIDAMDVQQ LSQIVPKVAV 600
 FYRASPRHKM KIKISLQRNG SVVAMTGDGV NDAVALKAAD IGVAMGQTGT DVCKEAADMI 660
 LVDDDFQITM SALESBGKIY NNINKFVRFQ LSTSIALLTL ISLATLMNEP NPLNAMQILW 720
 40 INIIMDGPFA QSLGVPEVDK DVIRKPPRNW KDSILTKNLI LKILVSSIII VCGTLFVFWR 780
 ELKINIVITPR DTTMTFTCFV PFDMENALSS RSQTESVFEI GLCSNRMPFY AVLGSIMGQL 840
 LVLYFPPLQK VFQTESLSIL DLLFLLGLTS SVCIVABIIK KVEREREKIQ KHVSTSSSF 900
 LEV

45 **A189 DNA SEQUENCE**
 Gene name: ESTs
 Unigene number: Hs.293185
 Probeset Accession #: N62096
 Nucleic Acid Accession #: N62096
 50 Coding sequence: 1-1284 (underlined sequences correspond to start and stop codons)

55
 60
 65
 70
 75
 1 11 21 31 41 51
 ATGGGCTACC AGAGGCAGGA GCCGTGTCATC CCGCCSCAGA GAGGATGCC TTATTCATG 60
 AAGCAAGCTG GGTTCCTCTT GGGAAATATT CTTTTATCTI GGGTTTCATA TGTACAGAC 120
 TTTTCTCTG TTTTATGAT AAAAGGAGGG GCCCTCTCTG GAACAGATAC CTACCACTCT 180
 TTGGTCAATA AAACCTTCGG CTTTCCAGGG TATCTGCTCC TCTCTGTTCT TCAGTTTGTG 240
 TATCTCTTTA TAGCATATGAT AGGTACAAAT ATAATAGCTG GAGATACTTT GAGCAAAAGT 300
 60 TTTCAAGAA TCCCAAGAGT TGATCCTGAA AACGTGTTTA TTGGTCGCCA CTTCATTATT 360
 GGACTTTCCA CAGTTACCTT TACTCTGCCT TTATCTCTGT ACCGAAATAT AGCAAAAGCT 420
 GGAAAGGTCT CCTCATCTC TACAGGTITA ACAACTCTGA TTCTTGGAA TGTAAATGGCA 480
 AGGSCAATTT CACTGGGTCC ACACATACCA AAAACAGAA AGCTTGGGT ATTTCRAAG 540
 CCCAATGCCA TCAAGCGGT GGGGTTATG TCTTTTGCAT TTATTTGCCA CCATAACTCC 600
 65 TTTCTAGTTT ACAGTTCTCT AGAAGAACCC ACAGTAGCTA AGTGGTCCCG CCTATCCAT 660
 ATGTCCATCG TGATTTCTGT ATTATCTGT ATATCTTTG CTACATGCG ATACTTGACA 720
 TTTACTGGCT TCACCCAAGG GGACTTATT GAAATTAAT GCAGAAATGA TGACCTGGTA 780
 ACATTTGGAA GATTTTGTTA TGGTGTCAT GTCAATTTGA CATACCTTAT GGAATGCTT 840
 GTGACAGAG AGGTAATTCG CAATGTGTTT TTGGTGGGA ATCTTTTCATC GGTTTTCCAC 900
 70 ATTGTGTAA CAGTGATGGT CATCACTGTA GCCACGCTTG TGTCAATGCT GATIGATTGC 960
 CTCGGGATAG TTCTAGAACT CAATGGTGTG CTCTGTGCAA CTCCCTCAT TTTTATCAT 1020
 CCATCAGCT GTTATCTGAA ACTGTCTGAA GAACCAAGGA CACTCTCCA TAAGATTATG 1080
 TCTTGTGTCA TGCCTCCAT TGGTGCTGTG GTGATGGTTT TTGGATTGCT CATGGCTATT 1140
 ACAAAATACT AAGACTGCAC CCAATGGGCG GAAATGTTCT ACTGCTTTC TGACAATTC 1200
 75 TCTCTCAGAA ATACTCAGA GTCTCATGTT CAGCAGACAA CACAACITTC TACTTTAAAT 1260
 ATTAGTATCT TTCAACTCGA GTAA

80 **A190 Protein sequence:**
 Gene name: ESTs
 Unigene number: Hs.293185
 Probeset Accession #: N62096
 Protein Accession #: none found
 Signal sequence: none found
 Transmembrane domains: 28-50, 66-88-112-134, 142-164, 217-239, 260-282, 298-320, 327-348, 359-381

Cellular Localization: plasma membrane

5
1
|
MGYQRQEPVI PPQGRGLPYSM KQAGFPLGIL LLFWVSIVTD FSLVLLIKGG ALSGTDITYQS 60
LVNKTGFPFG YLLLSVLQFL YPFIAMISYN IIAGDTLSKV FQRIQGVDPF NVPIGRHPFI 120
GLSTVITFLP LSLYRNIAKL GKVSLISTGL TTLILGIWMA RALSLGPHIP KTEDAWVPAK 180
PNAIQAVGVM SFAPICRHS PLVYSSLEP TVAKWRLIH MSIVISVPIC IFFATCGYLT 240
10 FTGFTQGDLE ENYCRNDLV TFGRFCYGVV VILTYPMECF VTREVIANVF EGGNLSVVFH 300
IVVTVMVITV ATLVSLLIDC LGIVLELNGV LCATPLLIYI PSACYLKLSE EPRTHSDKIM 360
SCVMLPIGAV VMVFGEVMAI TNTQDCTHQ EMFYCFPDNF SLTNTSESEV QQTITQLSTLN 420
ISIFQLE

15 A191 DNA SEQUENCE
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #: N62096
Nucleic Acid Accession #: N62096
20 Coding sequence: 1-1203 (underlined sequences correspond to start and stop codons)

25
1
|
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCGGT TTTCCCTTGT TTTATTGATA 60
AAAGGAGGGG CCTCTCTCTG AACAGATACC TACCAGTCTT TGGTCAATAA AACTTTCGGC 120
TTTCCAGGGT ATCTGCTCTC CTCTGTTCTT CAGTTTITGT ATCCTTTTAT AGCAATGATA 180
AGTTACAATA TAATAGCTGG AGATACTTGG AGCAAAGTIT TCAAAAGAAAT CCCAGGAGTT 240
GATCCIGAAA ACGTGTTTAT TGGTCGCGAC TTCATTATTG GACTTTCCAC AGTTACCTTT 300
30 ACTCTGCCCT TATCCTTGTA CGGAAATATA GCAAAGCTTG GAAAGGTCTC CCTCATCTCT 360
ACAGGTTTAA CAACCTCTGT TCTTGGAATT GTAATGGCAA GGGCAATTTC ACTGGGTCCA 420
CACATAACAA AAACAGAAGA CGCTTGGGTA TTTGCAAAGC CCAATGCCAT TCAAGCGGTC 480
GGGGTTATGT CTTTGGCATT TATTGGCCAC CATAACTCCT TCTTAGTTTA CAGTTCTCTA 540
GAAGAACCAC CAGTGTCTAA GTGGTCCCGC CTTATCCATA TGTCATCGT GATTTCGTA 600
35 TTTATCTGTA TATCTTTTGC TACATGTGGA TACTTGACAT TTAGTGGCTT CACCCAAGGG 660
GACTTATTTG AAAATTACTG CAGAAATGAT GACCTGGTAA CATTTGGGAG ATTTTGTAT 720
GGTGTCACTG TCATTTTGAC ATACCCATAG GAATGCTTTG TGACAAGAGA GGTAAATGGC 780
AATGTGTTTT TTGGTGGGAA TCTTTCATCG GTTTTCCACA TTGTGTAAAC AGTGATGGTC 840
ATCACTGAGT CCACGCTTGT GTCATTGCTG ATTGATTGCC TCGGGATAGT TCTAGAAGTC 900
40 AATGGTGTGC TCTGTGCAAC TCCCTCTATT TTTATCATTG CATCAGCTTG TTAATCGAAA 960
CTGTCTGAAG AACCAGGAC ACACCTCCAT AAGATTATGT CTGTGTCTAT GCTTCCCAT 1020
GGTGTCTGAG TGATGGTTTT TGGATTGCTC ATGGCTATTA CAAATACTCA AGACTGCAAC 1080
CATGGGCAAG AATGTTCTA CTGCTTCTCT GACAATTCTT CTCTCACAAA TACCTCAGAG 1140
45 TCTCATGTTT AGCAGACAAC ACACTTTCTT ACTTTAAATA TTAGTATCTT TCAACTCGAG 1200
TAA

50 A192 Protein sequence:
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #: N62096
Protein Accession #: none found
Signal sequence: 1-26
55 Transmembrane domains: 45-61, 92-108, 117-132, 191-207, 274-290, 297-313, 335-351
Cellular Localization: plasma membrane

60
1
|
MGYQRQEPVI PPOFSLVLLI KQALSGTDT YQSLVKNKTFG FPGYLLSVL QFLYFFIAMI 60
SYNIIAGDTL SKVQRIQGV DPENVFIGRH FIIGLSTVTF TLPLSLYRNI AKLGRVSLIS 120
TGLTLLILGI VMARISLGP HIFKTEDANV FAKPNAIQAV GVMSFAPICH HNSFLVYSSL 180
EEPTVAKWNR LIHMSIVISV FICIFPATOG YLFTGTGQQ DLFENYCEAD DLVTFGRFCY 240
65 GVTVILTYPM ECFVTREVI NVFFGGNLS VEHIVVTVMV ITVATLVSL IDCLGIVLEL 300
NGVLCATPLI FIIPSACYLK LSEPRTHSD KIMSCVMLPI GAVVNVFGFV MAITNTQDCT 360
HQQEMFYCFP DNPSTNTSE SHVQTTQLS TLNISIFQLE

70 A193 DNA SEQUENCE
Gene name: ESTs
Unigene number: Hs.293185
Probeset Accession #: N62096
Nucleic Acid Accession #: N62096
75 Coding sequence: 1-1140 (underlined sequences correspond to start and stop codons)

80
1
|
ATGGGCTACC AGAGGCAGGA GCCTGTCATC CCGCCGCGGT TCAATAAAAC TTTCCGCTTT 60
CCAGGGTATC TGCTCCTCTC TGTTCTTCAG TTTTGTATC CTTTATAGC AATGATAAGT 120
TACAATATAA TAGCTGGAGA TACTTTGAGC AAAGTTTTTC AAAGAATCCC AGGAGTTGAT 180
CCTGAAAACG TGTTTATTTG TCGCCACTTC ATTATGGAC TTTCCACAGT TACCTTTACT 240
CTGCCCTTTAT CCTTGTACCG AATATAGCA AAGCTTGGAA AGGTCTCCT CATCTTACA 300
GGTTTAACAA CTCTGATCTT TGGAAATGTA ATGGCAAGGG CAATTTCACT GGGTCCACAC 360
ATACCAAAAA CAGAAGAGC TTGGGTATTT GCAAAGCCCA ATGCCATTCA AGCGGTGCGG 420

GTTATGTCCTT TTGCAATTTAT TPGCCACCAT AACTCCTTCT TAGTTTACAG TTCTCTAGAA 480
 GAACCCACAG TAGCTAAGTG GTCCCGCCTT ATCCATATGT CCATCGTGAT TTCTGTATTT 540
 ATCTGTATAT TCCTTGCTAC ATGTGGATAC TTGACATTTA CTGGCTTCAC CCAAGGGGAC 600
 TTATTTGAAA ATTACTGCGA AAATGATGAC CTGGTAACAT TTGGAAGATT TTGTTATGCT 660
 GTCACCTGCA TTTTGACATA CCTATGGAA TGCTTTGTGA CAAGAGAGGT AATTGCCAAT 720
 GTGTTTTTTG GTGGGAATCT TTTCATCGGT TTCCACATG TTGTAACAGT GATGGTCATC 780
 ACTGTAGCCA CGCTTGTGTC ATTGCTGATT GATTGCTCG GGATAGTTCT AGAAGTCAAT 840
 GGTGTGCTCT GTGCAACTCC CCTCATTTTT ATCATTCAT CAGCCTGTTA TCTGAAACTG 900
 TCTGAAGAAC CAAGGACACA CTCCGATAAG ATTATGCTT GTGTCATGCT TCCCATTGGT 960
 GCTGTGGTGA TGTTTTTTGG ATTCGTCATG GCTATTACAA ATACTCAAGA CTGCACCCAT 1020
 GGGCAGGAAA TGTTCTACIG CTTTCTGAC AATTCTCTC TCACAAATAC CTCAGAGTCT 1080
 CATGTTCCAG AGACAACACA ACTTCTTACT TAAATATTA GTATCTTTCA ACTCGAGTAA

A194 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.293185
 Probeset Accession #: N62096
 Protein Accession #: none found
 Signal sequence: none found
 Transmembrane domains: 24-40, 70-86, 95-111, 171-186, 253-269, 276-292, 314-330
 Cellular Localization: plasma membrane

1 11 21 31 41 51
 NGYQRQSFVI PPQVKNKTFP PGVLLLSVLQ FLYPFIAMIS YNIIAGDTLS KVFQRIQVVD 60
 PENVFIRGRH IIGLSTVFTT LPLSLYRNIA KLKVSLSIST GLTTLILGIV MARAISLSPH 120
 IPKTEDAVWF AKPMIAIQAQ VMSFAFICHH NSFLVYSSLE EPTVAKWSRL IHMSIVISVF 180
 ICIFPATGGY LPTFGPTQSD LFENYCRND LVTGFRFCYG VTVILTYFMB CFVTREVIAN 240
 VFFGGLNLSV FHLVVVTVMVI TVATLVSLLI DCLGIVLELN GVLCAITPLIF IIPSACYLKL 300
 SEEPRTSHDK IMSCVMLEPIG AVVMVFGFVM AITNTQDCTH GQEMFYCFPD NPLSLNISES 360
 HVQQTQLSLT LNISIFQLE

A195 DNA SEQUENCE:

Gene name: ESTs
 Unigene number: Hs.293185
 Probeset Accession #: N62096
 Nucleic Acid Accession #: N62096
 Coding sequence: 1-1389 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 ATGGGCTACC AGAGGCAGGA GCCTGTCTAT CCGCGCAGGA GAGATTTAGA TGACAGAGAA 60
 ACCCTTGTTT CTGAACATGA GTATAAAGAG AAAACCTGTC AGTCTGCTGC TCCTTTTAAT 120
 GTTGCTCACT CGATTATAGG ATCTGGTATA ATAGGATTGC CTTATTCAAT GAAGCAAGCT 180
 GGGTTTCCYT TGGGAATATT GCCTTTATTC TGGGTTTCAT ATGTTACAGA CTTTTOCCYT 240
 GTTTTATTGA TAAAGGAGG GGCCTCTCTT GGAACAGATA CCTACCAATC TTGGGTCAAT 300
 AAAACCTTCG GCCTTCCAGG GTATCTGCTC CTCCTCATTC TTCACTTTT GTATCCTTTT 360
 ATAGCAATGA TAGGTTACAA TATAATAGCT GGAGATACTT TGAGCAAGT TTTTCAAGA 420
 ATCCGAGGAG TTGATCTCGA AAACGCTGTT ATTGGTCTGC ACTTCATTAT TGGACTTTCC 480
 ACAGTTACCT TTAATCTGCC TTTATCTCTG TACCAGAAAT TAGCAAGCT TGGAAAGGTC 540
 TCCCTCATCT CTACAGGTTT AACAACTCTG ATTCTTGGAA TTGTAATGGC AAGGGCAATT 600
 TCACTGGGTC CACACATACC AAAAACAGAA GACGCTTGGG TATTTGCAAA GCCCAATGCC 660
 ATTCAAGCGG TCGGGGTTAT GTCTTTTGCA TTTATTTGCC ACCATAACTC CTCTCTAGTT 720
 TACAGTTCTC TAGAAGAACC CACAGTAGCT AAGTGGTCCC GCCTATATCA TATGTCATC 780
 GTGATTTCTG TATTTATCTG TATATCTTT GCTACATGTC GATACTTGAC ATTACTTGGC 840
 TTCACCCAGG GGGACTTATT TGAATAATAC TGCAGAAATG ATGACCTGGT AACATTGGA 900
 AGATTITGTT ATGTTGTCTC TGTCATTTTG ACATACCCTA TGGAAATGCT TGTGACAAGA 960
 GAGGTAATG CCAATGCTT TTTTGGTGGG AATCTTTTAT CGGTTTTCCT CATTTGTGTA 1020
 ACAGTGAATG TCATCAGCT AGCCAGGCTT GTGTCTATGC TGATTGATTG CCTCGGAGTA 1080
 GTTCTAGAAC TCAATGGGCT GCCTCTGACA ACTCCCTCTA TTTTATCAT TCCATCAGCC 1140
 TGTATCTGA AACTGTCTGA AGRACCAAGG ACACACTCGG ATAAGATTAT GTCTTGTGTC 1200
 ATGCTTCCCA TTGTTGCTGT GGTGATGGTT TTTGGATTCT TCATGGCTAT TACAAATACT 1260
 CAAGACTGCA CCCATGGGCA GGAATGTTTC TACTGCTTTC CTGACAAATT CTCTCTCACA 1320
 AATACCTCAG AGTCTCATGT TCAGCAGACA ACACAACTTT CTACTTTAAA TATTAGTATC 1380
 TTTCAATGA

A196 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.293185
 Probeset Accession #: N62096
 Protein Accession #: none found
 Signal sequence: none found
 Transmembrane domains: 37-53, 66-82, 109-125, 155-172, 180-196, 255-271, 338-354, 361-377, 399-415
 Cellular Localization: plasma membrane

1 11 21 31 41 51
 NGYQRQSFVI PPQRDLDDRE TLVSEHEYKE KTCQSAALFN VVNSIIGSGI IGLPYSMKQA 60

GFPLGILLLF WVSXVTFDFSL VLLIKGGALS GTDTYQSLVN KTFGPPGYLL LSVLQFLYPP 120
 IAMISYNIIA GDTLSKVFOR IPGVDPENVF IGRHFTIGLS TVTFTLPLSL YRNIKLLGKV 180
 SLISTGLTTL ILGIWMARAI SLGPHLPKTE DAWVFAKENA IQAVGVMSFA FICHHNSFLV 240
 YSSLLEPTVA KWSRLIHMSI VISVPICIEF ATCGYLTFTG FTQGDLEFNY CRNDLIVTFG 300
 RFCYGVTVIL TYPMCECVTR EVIANVFFGG NLSSVFHIVV TVMVIIVATL VSLLDICLGI 360
 VLELNGVLCA TPLIFIIPSA CYLKLSEEPK THSDKIMSCV MLPIGAVVMV FGFVMAITNT 420
 QDCTHGQEMF YCFPDNFSLT NTSSEHVQQT TQLSTLNIISI FQ

A197 DNA SEQUENCE

Gene name: ESTs
 Unigene number: Hs.31608
 Probeset Accession #: H18836
 Nucleic Acid Accession #: NM_017636
 Coding sequence: 1-3501 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 1 | 11 | 21 | 31 | 41 | 51 |
 20 |
 ATGGAGGATG CCTTCGGGGC AGCCGTGGTG ACGGTGTGGG ACAGCGATGC ACACACCACG 60
 GAGAGGCCCA CCGATGCCCTA CCGAGAGCTG GACTTCACGG GGGCCGGCCG CAAGCACAGC 120
 AATTTCCTCC GGCCTCTCTGA CCGAAGCGAT CCAGCTGCAG TTATATAGTCT GGTTCACACG 180
 ACATGGGGCT TCCGTGCCCC GAACCTGGTG GTGTCACTGC TGGGGGGATC GGGGGGGCCC 240
 GTCTCCAGCA CCTGGCTGCA GGACCTGCTG CGTGTGGGCG TGGTGCAGGC TGCCACAGAG 300
 ACAGGAGCTT GATTTGTAC TGGGGGCTG CACACGGGCA TCGCCCGGCA TGTGGTGTG 360
 25 |
 GCTGTACGGG ACCATCAGAT GCCAGCACT GGGGGCACCA AGGTGGTGGC CATGGGTGTG 420
 GGGCCCTGGG GTGTGGTCCG GAATAGAGAC ACCCTCATCA ACCCCAGGGG CTGTTCCCTT 480
 GCGAGGTACC GGTGGCGCGG TGACCCGGAG GACGGGGTCC AGTTTCCCTT GGACTACAAC 540
 TACTCGGCTT TCTTCTCTGT GAGCGAGCGC ACACACGGCT GCCTGGGGGG CAGAGAACCG 600
 TTCCGCTGCT GCTTGGAGTC CTACATCTCA CAGCAGAGGA CGGGCGTGGG AGGGACTGGA 660
 30 |
 ATTGACATCC CTGTCTGTCT CTCTCTGATT GATGGTGTAT AGAAGATGTT GACGCGAATA 720
 GAGAACGCCA CCCAGGCTCA GCTCCCATGT CTCCTCGTGG CTGGCTCAGG GGGAGCTGGG 780
 GACTGCTTGG GCGAGACCTT GGAAGACACT CTGGCCCGAG GAGTGGGGG AGCCAGGCAA 840
 GCGAGAGCCC GAGATCGAAT CAGGCGTTTC TTCCCAAGG GGGACCTTGA GGTCTCTGAG 900
 GCGCAGGTGG AGGCGATTAT GACCCGGAG GAGCTCTGGA CAGTCTATTC TTCTGAGGAT 960
 35 |
 GGGTCTGAGG AATTCGAGAC CATAGTTTTC AAGGCCCTTG TGAAGGCTTG TGGGAGCTCG 1020
 GAGGCCCTAG CCTACCTGGA TGAGCTGCGT TTGGCTGTGG CTTGGAACCG CGTGGACATT 1080
 GCCCAGAGTG AACTCTTTGG GGGGGACATC CAATGGCGGT CCTTCCATCT CAGAGCTTCC 1140
 CTCATGGAGC CCCCTCTGAA TGACCGGCTT GAGTTCGTGC GCTTGTCTAT TTCCACGGC 1200
 CTCAGCTTGG GCGACTTCTT GACCCCGATG CGCTTGGCTT AACTCTACAG CGCGCGCCCC 1260
 40 |
 TCCAACTCGC TCATCCGCAA CCTTTTGGAC CAGGCGTCCC ACAGCGCAGG CACCAAAGCC 1320
 CCAGCCCTAA AAGGGGGAGC TGCGGAGCTC CGGCCCTCTG ACCTGGGGCA TGTGTGAGG 1380
 ATGCTCTCTG GGAAGACCTT GCGCCGAGG TACCTCTCCG GGGCGCGCTG GGAACCTCAC 1440
 CCAGGCCAGG GCTTCGGGGA GAGCATGTAT CTGCTCTCGG ACAAGGCCAC CTCGCGCTC 1500
 45 |
 TCGTGGATG AGTCCGCTCG CAGGCGCCCC TGGAGGAGC TGCTTCTTTG GGCAGTGTG 1560
 CTGAACAGGG CACAGATGGC CATGTACTTC TGGGAGATGG GTTCCAATGC AGTTTCTCTA 1620
 GCTCTTGGGG CCTGTTTGTCT GCTCCGGGTG ATGGCACGCC TGGAGCTTGA CGCTGAGGAG 1680
 GCAGACGCGA GGAAGAACCT GCGGTTCAAG TTGAGGGGGA TGGGCGTTGA CCTCTTTGGC 1740
 GAGTGTCTAT GCAGCACTGA GGTGAGGGCT GCGCGCTTCC TCCCTCGTGG CTGCGCGCTC 1800
 50 |
 TGGGGGGATG AGTCCGCTCG CAGCTTGGCC ATGCAAGCTG ACGCCCGTGC CTCTCTTGGC 1860
 CAGGATGGGG TACAGTCTCT GCTGACACAG AAGTGGTGGG GAGATATGGC CAGCACTACA 1920
 CCCATCTGGG CCGTGGTCTT GCGCTTCTTT TGCCCTCCAC TCATCTACAC CCGCTCATC 1980
 ACCCTCAGGA AATCAAGACA GAGGCCACA CCGAGGAGC TAGAGTTTGA CATGATAGT 2040
 GTCAATTAAT GGAAGGGGCC TGTGGGAGC GCGGACCCAG CCGAGAGAC GCGCTGGGG 2100
 55 |
 GTCCCGCGCC AGTCCGCTCG TCGGGTTCG TCGGGGGGCC GCTGCGGGGG GCGCGGTGC 2160
 CTAGCGCGCT GGTTCACCTT CTGGGGCGCG CGGTGACCA TCTTCATGG CAACGTGGTC 2220
 AGCTACCTGC TGTCTCTGCT GCTTTCTCTG CCGGTGCTGC TGGTGGATT CCAGCCGGCG 2280
 CCGCCCGGCT CCGTGGAGCT GCTGCTCTAT TTCTGGGCTT TCAAGCTGCT GTGCGAGGAA 2340
 CTGCGCCAGG GCGTGGCGCG AGGCGGGGGC AGCTCGCCA GCGGGGGGCC GGGGCTGGC 2400
 60 |
 CATGCTTAC TGAGCCAGCG CTGCGCTC TACCTCGCG ACAGCTGGAA CCAATGGGAC 2460
 CTAGTGGCTC TCACCTGCTT CCTCTGGGCT GTGGGCTGCC GCGTACCCCC GGGTTTGTAC 2520
 CACCTGGGCC GCACTGTCTT CTGCATCGAC TTCATGGTTT TCACGGTGGC GCTGCTTAC 2580
 ATCTTCACGG TCAACAACA CCGTGGGGCC AAGATCGTCA TCGTGAGCAA GATGATGAAG 2640
 65 |
 GAGTGTCTT TCTTCTCTT CTCTCTCGG GTGTGGCTGG TAGCCATGG CGTGGCCACG 2700
 GAGGGGCTTC TCGGGCCAGG GACAGTGAC TTCCCAAGTA TCTTGGCGCG GGTCTTCTAC 2760
 CGTCCCTACC TGCAGATCTT CCGGCAGATT CCGCAGAGG ACATGGAGCT GGCCTCATG 2820
 GAGCACAGCA ACTGCTCGTC GGAGCCCGGC TTCTGGGCAC ACCCTCTGG GCGCCAGGCG 2880
 GGCACCTGGG TCTCCAGTA TCCCAACTGG CTGGTGGTGC TGCTCTCTGT CATCTTCTG 2940
 70 |
 CTGTTGGCCA ACATCTGCT GGTCAACTTG CTCATTGCCA TGTTCAGTTA CACATTGGG 3000
 AAGATACAGG GCAACAGCGA TCTCTACTGG AAGCGCAGC GTTACCGCTT CATCGGGAA 3060
 TTCCACTCTC GCGCCCGGCT GCGCCCGGCC TTTATCTGTA TCTCCACTT GCGCTCTCT 3120
 CTCAGGCAAT TGTGAGGCG ACCCGGAGC CCGCAGCGT CCTTCCCGGC CCTGAGCAT 3180
 TTCCGGGTTT ACCTTTCTAA GGAAGCCGAG CGGAAGCTGC TAAAGTGGGA ATCGGTGCAT 3240
 AAGGAGAACT TTCTGCTGGC ACGCGCTAGG GACAAGCGGG AGAGCGACTC CGAGCGCTG 3300
 75 |
 AAGCGCAAGT CCGAGAGGT GCACTTGGCA CTGAAACAGC TGGGACACAT CCGGAGTAC 3360
 GAACAGCGCC TGAAGTGTCT GGAGCGGGAG GTCCAGCAGT GTAGCCCGCT CCTGGGGTGG 3420
 GTGGCCGAGG CCTGAGCGG CTCTGCTTGT CTGCCCCAG GTGGGCGGCC ACCCCCTGAC 3480
 CTGCTGGGT CCAAGACTG A

A198 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.31608
 Probeset Accession #: H18836
 Protein Accession #: none found

Signal sequence: none found
 Transmembrane domains: 214-230, 537-556, 642-662, 730-752, 760-782, 815-837, 842-864, 877-899, 973-995
 Cellular Localization: plasma membrane

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

```

1      11      21      31      41      51
|      |      |      |      |      |
MEDAFGAADV TVWDSDAHTT EKPTDAYGEL DFTGAGRKHS NPLRLSDRTD PAAVYSLVTR 60
TWGFRAPNLV VSVLGGSGGP VLQTLQDLR RRGIVRAAQS TGAWIVTGGL HTGIGRIVGV 120
AVRDHQMAST GGTKVAVMGV APNGVVRNRD TLINPKGSFP ARYRWRGDPE DGVQPFIDYN 180
YSAFLVDDG THGCLGGENR FRLRLSYIS QQTGIVGGTG IDIPVLLLI DGDEKMLTRI 240
ENATQALPC LLVAGSGGAA DCLAETLEDT LAPGSGGARQ GEARDRIIRF FPKGLDLEVLQ 300
AQVERIMTRK ELLTVYSSSD GSEEPETIVL KALVKACGSS EASAYLDELK LAVAWNRVDI 360
AQSELEFRGI QWRSEHLEAS LMDALLNDRP EFVRLLSHG LSLGHFLTPM RLAQLYSAAP 420
SNSLRNLID QASBSAGTKA PALKGGAAKL RPDVGVHVLK MLLGKMCAPR YPSGGAWDPH 480
PQGGFSGSMY LLSKDKATSP SLDAGLGOAP WSDLLWALL LNRAQMAMYF WEMGSNAVSE 540
ALGACLLLRV MARLEPDABE AARRKDLAFK FEGMGVDLFG ECVRSSEVRA ARILLRRCPH 600
WGDATCLQLA MQADARAFPA QDGVQSLLTQ KHWGDMASTT PIWALVLAFK CPFLIYTRLI 660
TFRKSEBEPF REELFEFMDG VINQEGFVGT ADPAEKTPLG VPRQSGRPGC CGRCGGRRC 720
LRRWPHFWGA PVTIIFMGVNV SYLLFLLLFH RVLIVDFQPA PPGSLLELLY FWAFTLLCEE 780
LRQGLSGGGG SLASGGGPGF HASLSQRLRL YLADSWNQCD LVALTCFLIG VGCRLTPGLY 840
ELGRVLCID FMVTVRLHLH IFTVNKQLGP KIVIVSKMKK DVFFFLFLG VMLVAYGVAT 900
EQLLRPRSDW FFSILRRVVF RPYLQIFQOI PQEDMDVALM ERSNCSSEPG FWAHPFQAQA 960
GTCBQYAMN LVVLLLVFL LVANILVNL LIAMFSYTFG KVQNSDLYN KAQRYRLIRE 1020
PHSRPALAPP FIVISHRLRL LRQLCRPRP PQPSSPALEH FRVYLSKEAB RKLLTWESVH 1080
KENFLARAR DKRESDSERL KRTSQKVDLA LKQLGHIREY EQRLKVLERE VQCSRVLGW 1140
VAFALRSAL LPFGGPPPPD LPGSKD

```

A199 DNA SEQUENCE

Gene name: ESTs
 Unigene number: Hs.40808
 Probeset Accession #: AA054237
 Nucleic Acid Accession #: AA054237
 Coding sequence: 1-894 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGAGAGCCGC GGGCGCTCGT CACGGCGCTC AGCCTGGGCC TCAGCCTGTG CTCCTCGGGG 60
CTGCTCGTCA CGGCCATCTT CACCGAACAC TGGTAGAGA CCGACCCCGG CGCCACAAG 120
GAGAGCTGCG AGCGCAGCCG CGCGGGCGCC GACCCCGCGG ACCAGAGAA CCGCTGATG 180
CGCTGTGCG ACCTGCGCTG CGCGGAGCTG CCCCCTGGTG GCGCGCGGCT GCTCCCGGGG 240
GGCCCGGGCG CGCGCGACCC CGAGTCTCTG CGCTCGCTCC TGGGCTCGCG CGGCTCGGAC 300
GCGGAGTGGG GCGCGCCCCC CTTCGCCACC TACTCGGGCC TCTGGAGGAA GTGCTACTTC 360
CTGGGCATCG ACCTGGGACAT CGACACCTCT ATCTGAAAG GTATTGCGCA GCGATGCACG 420
GCCATCAAGT ACCACTTTTC TCAGCCCATC CGCTTGCAG ACATTCTTTT TAAITTAACC 480
AAGACCATAC AGCAAGATGA GTGGCACCTG CTTCATTAA GAAGAATCAC TGCTGGCTTC 540
CTCGGCATGG CCGTAGCCGT CCTTCTCTGC GCGTGCATTG TGGCCACAGT CAGTTCTTTC 600
TGGGAGGAGA GCTTAGCCCA GCACGTGGCT GGAATCTCTT TCTCATGAC AGGGATATTT 660
TGACCATTTT CCTCTGTAC TTATGCCGCC AGTATCTGAT ATGATTGAA CCGGCTCCCA 720
AAGCTAATTT ATAGCTGCTC TCTGATGATG GAACATGGTT ACAGCTGGTC CATCTTTTTC 780
GCTGTGTGCA GTTAGGCTT TATTGTGGCA GCTGGAGGTC TCTGCATGCG TTATCCGTTT 840
ATTAGCCGGA CCAAGATTGC ACAGCTAAG TCTGGCAGAG ACTCCACGCT ATGA

```

A200 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.40808
 Probeset Accession #: AA054237
 Protein Accession #: none found
 Signal sequence: 1-18
 Transmembrane domains: 179-201, 209-231, 257-279
 Cellular Localization: plasma membrane

```

1      11      21      31      41      51
|      |      |      |      |      |
MEPRALVTAL SLGLSLCSLG LLVTAIFTDE WYETDPRRHK ESCERSRAGA DFPDQKNRLM 60
PLSHLPLRDS PPLGRRLLPF GPGRADPESW RSLGLGLD ABCGRPLFAT YSLWRKCYF 120
LGIDRDIDTL ILKQIAQRCT AIKYHFSQPI RLRLNIPNLIT KTIQDQEWHL LHLRRITAGF 180
LGMNAVALLC GCIVATVSPF WEBSLTHVA GLLFLMTGIF CTISLCTVAA SISYDLNRLP 240
KLIYSLPADV EHGYSWSIFC AMCSLGFIVA AGGLCIAYPF ISRTKIAQLK SGQDSTV

```

A201 DNA SEQUENCE

Gene name: CDAL4
 Unigene number: Hs.26813
 Probeset Accession #: N32912
 Nucleic Acid Accession #: NM_016570
 Coding sequence: 1-1134 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGAGGCGAC TGAATCGGAA AAAAAGCTTA AGTTTGGTAA AAGAGTTGGA TGCCCTTCOG 60
AAGGTTCTCT AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAATAGCA 120

```

20	<u>A202 Protein sequence:</u>	
	Gene name:	CDA14
	Unigene number:	Hs.26813
	Probeset Accession #:	N32912
25	Protein Accession #:	NP_057654
	Signal sequence:	none found
	Transmembrane domains:	none found
	Cellular Localization:	nuclear

40 A203 DNA SEQUENCE
Gene name: dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2)
Unigene number: Hs.44926
45 Probeset Accession #: S79876
Nucleic Acid Accession #: NM 001935.1
Coding sequence: 76-2301 (underlined sequences correspond to start and stop codons)

926

5
 10
 15
 20
 25

```

STATCCCGGT GGGAGTACTA TGACTCAGTG TACACAGAAC GTTACATGGG TCTCCCAACT 2100
CCAGAAGACA ACCITGACCA TTACAGAAAT TCAACAGTCA TGAGCAGAGC TGAAAAITTT 2160
AAACAAGTTG AGTACCTCCT TATTATGGA ACAGCAGATG ATACCGTTCA CTTTCAGCAG 2220
TCAGCTCAGA TCTCCAAGC CCTGGTCGAT GTTGGAGTGG ATTTCCAGGC AATGTGGTAT 2280
ACTGATGAAG ACCATGGGAT AGCTAGCAGC ACAGCACACC AACATATATA TACCCACATG 2340
AGCCACTTCA TAAACAATG TTTCTCTTTA CCTTAGCACC TCAAAATACC ATGCCATTTA 2400
AAGCTTATTA AAACCTATT TTGTTTTCAT TATCTCAAAA CTGCACGTGC AAGATGATGA 2460
TGATCTTTAA AATACACACT CAATCAAGA AACTTAAGGT TACCTTTGTT CCCAAATTTT 2520
ATACCTATCA TCTAAGTAG GGACTTCTGT CTTCACACA GATTATTACC TTACAGAGT 2580
TTGAATTATC CGGTGCGGTT TTATGTGTTA AAATCATTTT TGCATCAGCT GCTGAAACAA 2640
CAAAATAGGA TTGTTTTTAT GGAGGCTTTG CATAGATTCC CTGAGCAGGA TTTTANTCTT 2700
TTTCTAAGTG GACTGTGTTCA AATGTTGTTT TCTTCTTTAA AGGGATGGCA AGATGTGGGC 2760
AGTGATGTCA CTAGGGCAGG GACAGGATAA GAGGGATTAG GGAGAGAAGA TAGCAGGGCA 2820
TGGCTGGGAA CCCAAGTCCA AGCATAACAA CACGAGCAGG CTACTGTICAG CTCCCTCGG 2880
AGAAGAGCTG TTCACCAOGA GACTGGCACA GTTTTCTGAG AAGACTATT CAAACAGTCT 2940
CAGGAATCA AATATCGAAA GCACGTGACTT CTAAGTAAAC CACAGCAGTT GAAAGACTCC 3000
AAGAAATATG AAGCGAATCT GCGTCAACG CAGCCCCAG GTGCCAGTTA TGGCTATAGG 3060
TGCTACAAAA ACACAGCAAG GGTGATGGGA AAGCATTTGA AATGTGCTTT TAAAAAATAA 3120
TACTGATGTT CCTAGTGAHA GAGGCAGCTT GAAACTGAGA TGTGAACACA TCAGCTTACC 3180
CTGTTAAAG ATGAAATAT TGTATCACA AATCTTAAT TGAAGAGTC CTGCTATCAA 3240
TTTTTCTTAT TTCAITTCCT TGAGTGTCTT AATTAAAGA ATATTTTAA TCCCTTGGAC 3300
TCATTTTAA AATGGAACA TAAATACAA TGTATGTAT TATTATTCC ATTCTACATA 3360
CTATGGAATT TCTCCAGTC ATTTAATAAA TGTGCTTCA TTTTTC
  
```

A204 Protein sequence:

Gene name: dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2)
 Unigene number: Hs.44926
 Probeset Accession #: S79876
 Protein Accession #: NP_001926.1
 Signal sequence: none found
 Transmembrane domains: 6-28
 DPPIV_N term domain: 43-557
 Peptidase S9 domain: 558-635
 Cellular Localization: plasma membrane

40
 45
 50

```

1 11 21 31 41 51
| | | | |
MKTPNKILLG LLGAAALVTI ITVPVVLINR GIDBATADSR KTYTLTDYIK NYRLKLYSL 60
RWISDHEYLY KQENILVFN AEYGNSSVFL ENSTFDERGH SINDYSISPD GQFILLEINY 120
VKQWRHSYTA SYDIYDLNKR QLITEERIPN NTQNVWSPV GHKLAYVWNN DIYVKIEPNL 180
PSYRIITWTK EDIYNGITD WYVEEVFSA YSALWNSPFG TFLAYAQEND TEVPLIEYSF 240
YSDBSLQYFK TVRVYPKAG AVNPTVKFFV VNTDLSLSSVT NATSIQTAP ASMLIGDHYL 300
CDVTWATQER ISLQMLRRIQ NYSVMDICDY DESSGRWNCL VARQHIEGST TGVVGRFRPS 360
EPHFTLDGNS FYKLISNEEG YRHICYFOID KKDCTFITKG TWEVIGIEAL TSDYLYYISN 420
EYKGMPPGGRN LYKIQLDIYT KVTCLSCELN PERQYYSVB PSKEAKYYQL RCGSGPLPLY 480
TLHSSVNDKG LRVLIEDNSAL DKMLQNVQMP SKKLDIFILN ETKFWYQML PPHFDKSKKY 540
PLLDVYVAGP CSQKADTVFR LNWATYLAST ENLIIVASFDG RSGGYQGDKI MRAINRRLGT 600
FEVEDQIEAA RQFSKMGFVD NKRIATNGMS YGGYVTSNVL GSGGCVFKCG IAVAPVSRWE 660
YDSDVYTERY MGLPTFDENL DHYRNSVMS RAENPKQVEY LLHGATADDN VHFQBSAQIS 720
KALVDVGVDF QAMRYTDEDH GIASSTAHQR IYTHMSRPIK QCFSLP
  
```

A205 DNA SEQUENCE

Gene name: predicted exon
 Unigene number: none found
 Probeset Accession #: none found
 Nucleic Acid Accession #: none found
 Coding sequence: 1-261 (underlined sequences correspond to start and stop codons)

65
 70

```

1 11 21 31 41 51
| | | | |
ATGGCTCTGG CGAAGGTGAG GGAGCCAAAC GCAAAATGACA ATGCCATCAG AGTGACAAC 60
AGAAATGTGA TTAAGTGGCG TGCTAACCAAG TGTTCCTGTC ATGAGGCAGA AAGTGAATCC 120
AGAAACCCCTC AGGAGCTCTG GATGGGCGCTG CTCCTCTTGA TGGGGGTCTT AGAAGCATGT 180
GTGGAATAGA GGCCTCTGTC AGTCTGCTCC CTGAGAGATG ACAAGGAGCA GAGCCCCCAC 240
CAGCCACAC TGATGTCTA A
  
```

A206 Protein sequence:

Gene name: predicted exon
 Unigene number: none found
 Probeset Accession #: none found
 Protein Accession #: none found
 Signal sequence: none found
 Transmembrane domains: 46-68
 Cellular Localization: not determined

80

```

1 11 21 31 41 51
| | | | |
MALAKVREPQ ANDNAIRVDN RSVIKVRANQ CSLFEASES RNPQRLNMGL LLLMGVLEAC 60
VEMRPLSVWS LRDDKQSPH QPTLDV
  
```


A207 DNA SEQUENCE

Gene name: ESTs
 Unigene number: Hs.222886
 ProbeSet Accession #: AI672225
 Nucleic Acid Accession #: none found
 Coding sequence: 1-462 (underlined sequences correspond to start and stop codons)

```

10      1      11      21      31      41      51
      |      |      |      |      |      |
      ATGCCAAATG CTGAGTTAGA AGCAAAGAGC CTTGGAAAGCA GTAATGTTT AAAAACTGCT 60
      CTCACTCTG CTGATGTTG TGGATCAGCA AATATAGTCA GCCCTCTACT TGAGCAAAT 120
      ATTGATGAT CTCTCAAGA TCTGGACAGA CGGCCAGAGA GTATGCTGTT TCTAGTCATC 180
      ATCATGTGA CCAGTTTGT GGAAGACAAT CTTTCCATGG GCTGGGGGAA GCTAGAGAT 240
      TTTATGGCTA TTGAAGAAGA AATGAAGAAG CACGGAAAGTA CTCATGTGGG ATTCCAGAA 300
      AACCTGACTA ATGGTGCCGC TGCTGGCAAT GGTGATGATG GATTAATTCC TCCAAGGAAG 360
      AGCAGAACAC CTGAAGGCCA GCAATTTCT GACACTGAGA ATGAAGAGTA TCACAGGTTT 420
      GTCAAAGATC AGATAGTTGT AGATATGCGG CGTTATTCTT GA
  
```

A208 Protein sequence:

Gene name: ESTs
 Unigene number: Hs.222886
 ProbeSet Accession #: AI672225
 Protein Accession #: none found
 Signal sequence: none found
 Transmembrane domains: 16-38
 Cellular Localization: not determined

```

30      1      11      21      31      41      51
      |      |      |      |      |      |
      MPNAELRAKS LGSSKCLKTA LILAVCCGSA NIVSPLEQIN LDVSSQQLDR RPESMLFLVI 60
      IMWTSFVEDN LSMGWGKLED FMAISEEMKK HGSTHVGFPR NLITGAAGAGN GDDGLIPPRK 120
      SRTPESQQFP DTENEZYHRF VKDQIVVDMR RYF
  
```

A209 DNA SEQUENCE

Gene name: odz (odd Oz/ten-m, Drosophila) homolog 1
 Unigene number: Hs.23796
 ProbeSet Accession #: NM_014253
 Nucleic Acid Accession #: NM_014253
 Coding sequence: 65-8242 (underlined sequences correspond to start and stop codons)

```

45      1      11      21      31      41      51
      |      |      |      |      |      |
      GACTGCTTGC ATTAAGGAC TTCTCATCC TTTTTCATCAT GAAACTGAGC TTGCTTAATC 60
      AGAGATGGAG CAAACTGACT GCAAACCCCTA CCAGCCTCTA CCAAAAGTCA AGCATGAAT 120
      GGATCTAGCT TACACCAATT CTCTGATGA GAGTGAAGAT GGAAGAAAAC CAAGACAGTC 180
      ATCAACTGCC AGGGAGACCC TGCAAGAGTA TAACCAAGAG CTGAGGATGA ATTACAATAG 240
      CCAAGATAGA AAGAGGAAGA AAGTAGAATA ATCTACTCAA GAGATGGAAAT TCTGTGAAC 300
      CTCTCAACT CTGTGCTCTG GCTACCAAAC AGACATGCAC AGCGTTTCTC GGCATGGCTA 360
      CCAGCTAGAG ATGGGATCTG ATGTGGACAC AGAGACAGAA GGTGCTGCCT CACCTGACCA 420
      TGCACATAAG ATGGGATCTG GGGGAATGAA ATCAGAGCAT AGTTCTCTTT TGTCCAGCG 480
      GGCCAACTCT GCATTATCTT TGACTGACAC TGACCATGAA AGGAAGTCTG ATGGGGAAA 540
      TGGTTTCAA TTCTCTCTG TTTGTGTGA CATGGAGGCT CAAGCTGGGT CTACTCAAGA 600
      TGTGCAGAGC AGGCCACACA ACCAGTTTAC CTTCAGACCC CTCCACCGGC CACCTCGGOC 660
      TCTCTATGCC TGCACTCTGT CCAGGAAGCC ACCCCCTGCA GCGGACTCTC TTCAGAGGAG 720
      ATCAATGACT ACCCGAGCC AGCCCAAGCC AGCTGCTCCA GCTCCCCCA CCAGCAAGCA 780
      GGATTCAGTC CATCTGCATA ACAGCTGGGT OCTGAACAGC AACATAACAT TGGAGACGAG 840
      GCATTCCCTG TTCAACATG GATCTGGTTC CTCTGCGATC TTCAGTGCAG CCAGTCAAG 900
      CTACCTCTCTG ACATOCATA CCGTGTACTC GCGCCCTCCC AGGCTCTTTC CTCGAAGCAC 960
      CTTTTCGGA CTTCTCTTTA CCTTTAACA ACCTTACAG TGCTGCAACT GGAAGTGCAC 1020
      AGCATTGAGC GCACTGCAA TCACAGTGAC TTTGGCCTTG TTAAGTGCCT ATGTGATTGC 1080
      AGTGATTGTT TTGGCCTGTA CTGTCAGTT GCAACCAATT GAAGGAGAGC TGTATGCAA 1140
      TGGAGTTAGC AAGGGAACA GGGGGACCGA GTCCATGGAC ACTACTACT CTCCAATTGG 1200
      AGGAAAAGTT TCTGATAAAT CAGAGAAAA AGTGTTTCAG AAGGGACGGG CGATAGACAC 1260
      TGGAGAAAGT GACATGGTG CACAGGTCAT GCAGACCAT CCACCTGGIT TATTCTGGCG 1320
      TTTCCAGATT ACTATCCAAC ATCCAATATA TCTGAAGTTC AATATTTCIT TAGCCAAGGA 1380
      CTCTCTGCTG GGAATTTATG GCAGAAGAAA CATTCACCT ACACATACIT AGTTTGATT 1440
      TGTAAAACTA ATGGATGGCA AACAGCTGGT CAAGCAGGAC TCCAAGGGCT CTGATGATAC 1500
      ACAGCACTCC CCTCGGAACC TGATCTTAAC TTGCTTCAG GAGACAGGTT TCATAGAGTA 1560
      TATGGATCAA GGACCTTGGT ATCTGGCGTT TTACAATGAT GGAAAAAAGA TGGAGCAAGT 1620
      AATCGTGTTA ACTACAGCAA TGAATAATAT GGATGACTGT TCACCAATT GCAATGGAAA 1680
      TGGAGAGTGT ATCTCTGGCC ATTGTCTATT TTTCCAGGA TTCTTGGAC CTGACTGTGC 1740
      TAGAGATTCC TCCCTGTGTC TGTGTGGTGG GAATGGAGAA TACAGAAAAG GACACTGTGT 1800
      CTGCGGCGAT GGCTGGAAGG GGCCAGAGTG TGAGCTTCCG GAAGAACAAAT GCATTGATCC 1860
      AACATGCTTT GGCCACGGCA CCTGTCATCAT GGGAGTCTGC ATCTGTGTGC CAGGATACAA 1920
      AGGAGAAATA TGCCAGGAAG AGGACTGCGT AGACCCAAAT TGTTCCAACC ATGGCATCTG 1980
      TGTAAAGAGA GAATGTCACT GTTCTACTGG CTGGGGAGGA GTTAACTGTG AAACACCACT 2040
      TCTGTATGT CAGAGACAGT GCTCAGGACA OGGAACCTTT CTCTGAGAGC CTGAGATATG 2100
      CAGCTGTGAT CCCAAGTGA CAGGATCTGA CTGCTCAACA GAGCTGTGTA CCATGGAGTC 2160
      TGGTAGCCAT GGAGTCTGCT CAAGAGGAAT TTGCCAGTGT GAAGAGGGCT GGTAGGAGCC 2220
      AACATGTGAG GAACGCTCCT GTCAATCTCA TTGTACTGAG CATGGCCAAT GCAAGATGG 2280
  
```

	AAAAATGTGAG	TGTAGECCCTG	GATGGGAGGG	CGACCACTGC	ACAATTGCTC	ACTACTTAGA	2340
	TGCTGTCCGA	GATGGCTGCC	CAGGCTCTG	CTTTGGAAAT	GGACGATGTA	CCCTGGATCA	2400
	AAATGGTTGG	CACGTGTGTGT	GTGAGGTGGG	TGGAGTGGG	ACAGGCTGCA	ATGTTGTCTAT	2460
5	GGAAATGCTT	TGTGGAGATA	ACTTGGACAA	TGATGGAGAT	GTTTAAACCG	ACTGTGTGGA	2520
	TCCTGACTGT	TGTCACACAA	GCAACTGTTA	TATAAGTCTT	CTCTGCCAGG	GCTCACCAGA	2580
	TCCTCTTGAC	CTCATTACAG	AAAGCCAAAC	TCTCTTCTCT	CAGCACACTT	CAAGACTTTT	2640
	TTATGATGGA	ATCAAAATCC	TCAITGGCAA	GGACAGTACT	CATGTCAATC	CTCCTGAGGT	2700
	GTCAATTGAC	AGCAGGCGTG	CCTGTGTGAT	TCGAGGCCAA	GTGGTGGCCA	TAGATGGAAAC	2760
10	TCCTCTAGTG	GGAGTGAATG	TCAGTTTCTT	GCACCCACAGT	GATTATGGGT	TTACCATCAG	2820
	COGGCAAGAT	GGAGGCTTTG	ACCTCGTGGC	CATCGGTGGC	ATCTCTGTCA	TCCTTAATCTT	2880
	CGACCGATCC	CCCTTCTCTG	CTGAGAAGAG	AACACTCTGG	TTGCCCTTGA	ATCAGTTTAT	2940
	TGTGGTAGAG	AAAGTCACCA	TGCAGAGAGT	TGTATCAGAC	CCGCCATCCT	GCGATATCTC	3000
	CAACTTTATC	AGCCCAAAAC	CTATTGTGCT	TCCTTCACCG	CTCACATCAT	TTGGAGGGTC	3060
15	CTGTCCAGAG	AGGGGAACCTA	TTGTTCCGGA	GCITGCAGGT	GTACAGGAGG	AAATTTCCCAT	3120
	TCCTTCAGC	TTTGTGAGGC	TGAGTTACCT	GAGCAGCCGC	ACCCCTGGGT	ATAAAACCTT	3180
	GCTACCGATC	CTTCTGACAC	ATTCAACGAT	TCCCGTAGGC	ATGATAAAAG	TACACCTCAC	3240
	AGTAGCTGTG	GAGGGGCGAC	TCACACAGAA	GTGGTTTCCC	GCGCAATTA	ATCTTGTCTA	3300
	CACATTTGCT	TGGAACAGAA	CGATATCTTA	TGGACAGAA	GTTTGGGGCC	TGGCAGAGGC	3360
20	TTTGTATATC	GTGGGATATG	AATATGAAC	GTGCCCTGAC	TTTATTCTCT	GGGAGCAAG	3420
	GACAGTGGTT	TTTACAGGTT	TTGAGATGGA	TGCTTCTAAC	CTAGGAGACT	GCTCTTTGAA	3480
	TAGCATCAC	ATTCTGACAC	CTCAAGTGG	AATCATACAT	AAAGGGAATG	GAGAAAAATAT	3540
	GTTCATTCCC	CAGCAGCCCC	CAGTCATATC	AACATAATG	GGTAATGGAC	ACCAAGGAG	3600
	TGTAGCCCTG	ACCAACTGCA	ATGGCCGAGC	CCACAACAC	AAACTCTTTC	CTCCTGTCTC	3660
25	CTTAGCTTCT	GGCCCTGATG	GCAGTGTGTA	TGTTGGGAC	TTCAATTTTG	TAGGAGAAAT	3720
	ATTTCCCTCG	GGAAACTCCG	TTAGTATTTT	GGAATTAAGC	ACAAGTCTTG	CTCAAAATA	3780
	CTATCTGGCT	ATGGACCTTG	TGTCGTAATC	ACTCTATCTA	TCAGACACCA	ATACTCGCAA	3840
	AGTCTACAAG	TGTAAATCTC	TTGTGGAGAC	GAAAGATCTG	TCCAAGAAIT	TTGAAGTGGT	3900
	GGCAGGAATC	GGTGATCAGT	GCCTTCCCTT	TGACAGAGT	CATTGTGGAG	ATGGTGGGAG	3960
30	AGCATCCGAA	GCTTCACATG	ATAGCCCTCG	AGGCATCACA	GTTGATAGGC	ATGGATTTAT	4020
	TTACTTTGTG	GAGGAGCTTA	TGATTCGCAA	AATTGATGAG	AATGCTGTGA	TCACAACTGT	4080
	AATCGGCTCA	AATGGTCTGA	CTTCCACACA	ACCACAGAGC	TGTGACTCAG	GAATGGACAT	4140
	CACTCAGGTG	CGATTAGGTG	GGCCAAACAG	CCTTGCACTA	AATCCTATGG	ACAATTCATT	4200
	GTATGCTTTG	GATAACAACA	TTGTGCTGCA	AATTTCTGAG	AACAGGCGTG	TTCTGATCAT	4260
35	CGCAGGACCG	CCCATTCACT	GCCAGGTGCC	AGGCATCGAT	CATTTCCTGG	TCAGCAAGGT	4320
	AGCAATTAC	TCCACTCTAG	AGTCAGCGAG	GGCCATCAGT	GTCTCCACCA	GCGGGCTGCT	4380
	CTTCTAGCTG	GAAACAGACG	AGAGGAAAGT	AAACCGCATT	CAGCAAGTAA	CCACCAATGG	4440
	GGAGATCTAC	ATCATCTGCTG	GTGCCCCCAC	TGACTGTGAC	TGCANAATTG	ATCCAAACTG	4500
	TGACTGTTTT	TCAGGTGATG	GTGGCTATGC	CAAAGATGCA	AAGATGAAG	CCCTTCTCTC	4560
40	CTTAGCAGTG	TGCGCTGATG	GAACCTCTTA	TGTGGCAGAC	CTCGGAATG	TTGGAATTTG	4620
	TACCATCAGC	AGGAACCAAG	CCCACTGAA	TGACATGAAC	ATTATGAGA	TTGCTTCACC	4680
	CGCTGATCAG	GAACTGTACC	AGTTCACTGT	AAATGGAACC	CACCTACACA	CCCTGAACCT	4740
	GATAACAGG	GACTATGTTT	ATACTTCAC	CTACAATTCT	GAAAGTGAAT	TGGGCGCGAT	4800
	TACCGCAGC	AATGGCAATT	CAGTGCACAT	TCGCCCTGAT	GCAGGCGGAA	TGCCGCTATG	4860
45	GCTTGTGTTG	CCTGGCGGAC	AAGTATACCT	GCTGACTATA	AGCAGCAATG	GAGTCCCTGA	4920
	AAGAGTGTCA	GCCCAAGGCT	ATAATCCGGC	CTTAATGACC	TATCCAGGAA	ACACAGGGAT	4980
	TCGTGGCTACC	AAAGATTAAG	AAATGGATG	GACAACGTTT	TATGAGTATG	ACCCCGAGGG	5040
	ACACCTGAC	AATGCAAGT	TCCCACTGG	AGAGGTGAGC	AGCTTCCACA	GTGACCTGGA	5100
	GAAGCTGACA	AAAGTGGAGC	TAGATACCTC	CAACCGTGAA	AATGTCTCTA	TGTCAACCAA	5160
50	CTTGACGGCA	ACTAGTACCA	TATATATTTT	AAAACAAGAA	AATACTCAAA	GTACCTTATG	5220
	GGTGAATCCA	GATGCTTCCC	TGCGTGTAC	TTTTGCCAGC	GGGATGGAGA	TGCGCTCAG	5280
	CTCAGAGCCC	CACATCTCTG	CAGGGGAGT	CAACCTTACC	CTGGGCAAAAT	GCAACATCTC	5340
	ATTGCCCGGA	GAGCACAATG	CAACCTCAT	CGAGTGGCGG	CAGAGGAAGG	AGCAAAACAA	5400
	AGGCAATGTT	TCGGCTTTTG	AAAGGAGGCT	GAGGGCCAC	AACAGAAACC	TACTCTCCAT	5460
55	AGATTTTATG	CATATAACCC	GCACAGGAAA	GATCTATGAT	GHCCATCGAA	AATTCACCCCT	5520
	TOGAATTTCT	TATGACAGCA	CTGGGCGACC	CACTCTGTGG	TCTCCTGTAA	CGAGATATAA	5580
	TGAAGTGAAC	ATCACATATT	CACCTTCCGG	AATGGTGAGC	TTTATTCAAA	GAGGAACGTT	5640
	GAAAGAAATA	ATGGAATATG	ACCAGAGTGG	GAAATTAATT	TCAAGAACCT	GGGCTGATGG	5700
	GAAATTTTGG	AGCTATACCT	ACTTAGAAAA	ATCTGTGATG	CTTCTCCTAC	ACAGCCAGCG	5760
60	GCGTATACAT	TTTGAGTATG	ACCAATCAGA	TTGCCCTGCTG	TCAGTTACCA	TGCCGTAGCAT	5820
	GGTGGGCCAC	AGCTTACAAA	CCATGCTTTC	AGTGGGCTAC	TACCGTAATA	TCTACACCCC	5880
	ACCAGACAGT	AGCACTTCTT	TTATCCAGGA	CTATAGTCTA	GATGGCGGAT	TGCTACAGAC	5940
	CCCTGATCTG	GGGACAGGGC	GCAGAGTCTT	ATACAGTAC	ACCAAGCAAG	CAAGGCTTTC	6000
	TGAGGTCTTC	TATGATACCA	CTCAGGTCTC	ATTAACATAT	GAAAGTCTT	CTGGAGTGAT	6060
65	TAGACAAATA	CAOCTGATGC	ATGACGGATT	CATCTGCACA	ATCAGATACA	GGCAACAGG	6120
	ACCTCTTATT	GGAGCGCCAGA	TTTTAGATT	CAGTGAAGAA	GGCCTTGTGA	ATGCAAGGTT	6180
	CGACTACAGC	TACAACAATT	TCOGAGTCAC	AAGCATGCAA	GCTGTAAATCA	ATGAAAACCC	6240
	TTTGCCTATA	GATCTTTACC	GATATGTTGA	TGCTCTGGC	AGAACAGAGC	AGTTTGGAAA	6300
	ATTGAGTGA	ATTAAATTACG	ATTAAATCA	GGTCTAACT	ACTACAGTGA	TGAACACAC	6360
70	CAAAATCTTC	AGTGCCAAATG	GACAAGTCTAT	TGAAGTCCAA	TATGAATCTC	TAAAGGCAAT	6420
	TGCGTACTGG	ATGACCAATC	AATATGATAA	TGTGGGCGGA	CAATGGTAATA	TGTGCATAAG	6480
	GGTAGAGTGA	GATGCCAATA	TAAACAGGTA	CTTCTATGAA	TACGATGCTG	ATGGGCAACT	6540
	TCAGACTGTT	TCGTGAATG	ACAAAACCCA	GTGGCGTTAT	AGTTACGATC	TGAATGGAGA	6600
	CATCAACCTC	TTAAGCCATG	GGAGAGTGC	TGCTCTTACT	CTCTCCGAT	ATGACCTCCG	6660
75	AGACCGCATC	ACCAGATTAG	GAGAAATTC	GTATAAAATG	GATGAAGATG	GCTTTCTGAG	6720
	GCAGAGGGGA	AAATGATATT	TTGAATATAA	TTCTAATGGC	CTGCTGCAGA	AAGCCTACAA	6780
	TAAAGCTTCT	GGCTGGAGTG	TGCAGTATTA	CTATGATGGG	CTTGGGCGAC	GTGTCCGAG	6840
	TAGTCCCTG	CTAGGGCAGC	ACCTTCAGTT	CTTGTGAGC	GCGACCGCGA	ACCCATAAG	6900
	AGTTACTCAT	TTGTACAACT	ACCAAGCTC	GGAGATTACA	TCTCTGTATT	ATGATCTCCA	6960
80	AGGTCACTTT	ATTGCCATGG	AGTTAAGCAG	TGCTGAAGAA	TATATGTAG	CTGTGTATAA	7020
	TACAGTACC	ATGACCTGCTG	TGTTACGACG	CCGAGGTCAG	GTCAATAAGG	AGATACATA	7080
	CACACCTTAT	GTCGATATCT	ATCATGACAC	ITACCTGAC	TTTCAGGTCA	TAAATGGTTT	7140
	TCATGGAGGA	CTCTATGATT	TCCTTACTAA	ATTAGTGAC	CTGGGGCAAA	GGGATATGA	7200
	TGTTGTGTCT	GGCAGATGGA	CAACGGCCTA	TCATCACATA	TGGAAACAGT	TGAACCTCCT	7260
	TCCTAAACCA	TTCAACCTCT	ACTCCTTTGA	AAATAACTAC	CCAGTTGGCA	AAATTCAGA	7320

	TGTTGCAAG	TATACCACAG	ACATCAGAAG	TIGGTTGGAG	CTATTGGT	TCCATTACA	7380
	CAATGTACTA	CCTGGATTTC	CCAAACCTGA	ATTAGAAAT	TTAGAAATTA	CTTACGAGCT	7440
	TCTACGGCTT	CAGACAAATA	CTCAAGAGTG	GGATCCTGGA	AAGACTATCC	TGGGCATTCA	7500
5	GGTGAACCTC	CAGAAACAGC	TCAGGAATT	CAITTCCTTG	GACCAACTAC	CTATGACTCC	7560
	CCGATACAT	GATGGACGGT	GCCTTGAAGC	AGGGAAGCAA	CCAAGGTTTG	CTGCTGTCCC	7620
	TTCTGTTTTT	GGGAAGGTA	TAAAATTTCG	CATCAAGGAT	GGCATACTAA	CAGCTGATAT	7680
	TATAGGAGTA	GCCAAATGAG	ATAGCAGGCG	GCTTGCTGCC	ATTCTCAATA	ATGCCCATTA	7740
	CCTGGAAAC	CTACATTTTA	CCATAGAGGG	GAGGGACACT	CACTACTTCA	TTAAGCTTGG	7800
10	GTCTCTGGAG	GAAGACCTGG	TGCTCATCGG	TAACACTGGG	GGGAGGCGGA	TTCTGAGAA	7860
	TGGTGTCAT	GTCACTGTGT	CCCAGATGAC	TTCTCTGTTG	AATGGGAGGA	CTAGACGGTT	7920
	TGCAGATATT	CAGCTCCAGC	ATGGAGCCCT	GTGCTTCAAC	ATCCGGTATG	GGACAACGT	7980
	CGAAGAGGAA	AAGAATCAAG	TGTTGGAGAT	TGCCAGACAG	CGCGCAGTGG	CCAGGCCCTG	8040
	GACTAAGGAA	CAGAAGAGGC	TGCAAGAGGG	GGAAGAGGGG	ATTAGGGCAT	GGACAGAAGG	8100
	GGAAAAGCAG	CAGCTTTTGA	GCACTGGGCG	GGTACAAGGT	TACGATGGGT	ATTTTGTITT	8160
15	GTCTGTTGAG	CAGTATTTAG	AACITTTCTGA	CAGTGCCAA	AATATTCACT	TTATGAGACA	8220
	GAGCGAAATA	GGCAGGAGGT	AACAAAATA	TCTCTGCCTT	TGCGTCAACA	AAGACTGCCT	8280
	GTTTTAAAAA	TAAATAATGG	TTTATTGTAT	TGGTTTTCTA	GATCAGAACT	CTGTATATGT	8340
	AATATGGAG	GAATAACATA	TCCAACCTGC	TTCAATGTG	ACGGAAGATG	GTATTTTAAT	8400
	ATTGTTTGT	TAACTCTTT	AAGAATGAC	AGAGATTTT	AGTTCCTG	TGGCAGTATT	8460
20	CABAATAACA	CAGTAGAAC	TCAAACAGCT	AAAAACAGTT	TTCBGAAGC	ACCACITTTCA	8520
	ATTTCGCGAG	CCATGCATAT	GTTCCAATAT	CCAGAAAGAA	CCCAAGGTTT	TCATCTCTTA	8580
	TTGTGAGGAA	AAGAATCAAG	CTTAACCTGT	GGCAGAACTT	ACGGGCTATT	TGAATAGGTG	8640
	GTGCAATAGT	ATCTGAAAT	TGCCITTCGA	AAGACTGCCA	GCCTTTTGAC	GTTTTCCAGA	8700
	TCTGTATAG	CAGCTTTTAA	AACAGGTGTA	AAATGTCTTC	AGCCACCATC	TCTTAGAGTG	8760
25	AGGACCAAT	TGCCCTTCTT	TCTGTATTAT	TCCTCCTTGC	TTGTTAAAGT	AAATGCCATA	8820
	TTGTTGTGCT	GTGTTTTGGC	GTGTGGTGGC	TGGGTTCTGT	CTACCATGCT	TCCCTGTGGG	8880
	TGTGGTAACC	AGACTGTATA	GCGCTATTAT	GCTCGTGTGT	ACATGATACC	AAAGCAGCTG	8940
	GCCAGCGTGA	CCCTCTCTAC	ACGACCTGTT	TTGACTCAAT	TTTTTACTAA	AAGTTGTTCA	9000
	GCTGTATTGG	TATCATGTAA	ACATAGCTTT	TATTAACCTG	GGTAGGAATT	TCTCATTTAT	9060
30	ATATAGGATG	TGTTTTGGTC	ATAGTTTCAC	ATTAGTGATT	CAGTATCTAT	ACACTGACCC	9120
	AATGGTTTTG	TGCACATGAA	CGGTAATTTA	CTTAAAGATA	TGATTTCTGGT	ACAAAAACAA	9180
	ACAAAGGCTT	TAGCAGGCAT	ACGTGTCTGG	GATGCCGATA	CATACATTA	CTACTACTGC	9240
	AGAAATTCAT	AAGAGCCAAA	ACCTTAAAAA	AATAGACCTG	GTACTTAAAT	GAAAGTACTA	9300
	AAGGGAAGAC	CAGACCAAC	ATCAGACGAG	TGCTGOCAC	ATTGTTTCAG	CCCACTAGA	9360
35	TTATCTTTC	AAATGTACAA	TTCTGTATTG	AACATCTCCC	AGCCATCTTC	AGGAAATCGA	9420
	ATCAAGTAAA	TCCTTTCCAA	CCGAAAACAT	TTCAACTAAC	TATAGAGAGG	CAGACTCAAT	9480
	TTTACTAAAA	TAAATATATC	AGTTAGTTAT	TTTGTCTCTC	CGTACTTACC	CATTTATCTT	9540
	TATTTAATCG	TCTCTACTGC	CTAGGAAAT	AACATTTTTC	CAGGACGGGT	TATTTGTTCT	9600
	GCGATCATTT	AAATTTTGA	GAAAGGTGAG	GATTAGTGTT	AATATCAGCT	GCACTTTCTC	9660
40	AATCTTAGG	AATCCTGCGAG	TAAACACAGC	CCCTTGGTGA	GCTGGAAGAT	TTGTGCCGAG	9720
	TGACAAAGAG	ATAGTTTGTA	AAATGCTGTG	TAAITGTAA	TAACACAAA	TGAAATATA	9780
	TGACAGCACA	ATGTGGCCCG	TAGAAAATTC	CCCTGAGCCA	GCTTCTGCAC	TTTCATCAC	9840
	GAATCTGAAC	ATTGTCTATG	TCTGAGGCA	AATTTATGAT	GGAATGTAG	TTTGGATTCT	9900
	TTCCAGATGC	TACCTAAATG	CAGTGTGGGG	TCATGCTT	GCTTTGCBAT	GACATTTCT	9960
45	TTGAAATAT	GCAAAATCAT	AAGCTCATGT	TAAGGTTTTT	CAAGASTCTG	CCTCCTACTA	10020
	CACAAAGGAA	AGCAAGGGAA	AGGAAATGAC	CCTGGCAAC	AGTAGGGAG	GGTGTATTCA	10080
	AACATTTCT	TTTCAAAACC	TTGGGGTTAG	AATACCACTT	ACACATGTAT	TCTGAGAGAC	10140
	AGAAATCATG	AGGAATCAT	CTCTCTTTAT	AACCTGAAAC	ACACCAGCTT	GATATATTGC	10200
	TAAATCCATA	TAAATCATTA	TTATTTGGTT	TTTTCTGAAT	CAGGCTCTGA	TTAATGGTAC	10260
50	AGTATTTAT	CAGAAATGAA	TTCTAAATTT	ACTAACAAAC	TTGTTGAAAA	TTTGAATACC	10320
	TCACACCCAA	CTTAAATAGT	GACCTTAAGT	TCCTAGAAC	TCTGATGTTT	TTTTAAATTA	10380
	ATGGAATAT	AATTTGTGAA	CTGTATATAG	AGAGTGCAAT	CATAAATGTG	ATTATGTATT	10440
	TTATCACAAT	TCCAAATAGT	CAATATTAGA	GTCTATTTTG	CTTATATTTT	AAGCAATTAT	10500
	ACGTTTGTGC	ATATCATTTA	TGAATGATCA	TTTTCAAAC	GCTTTAAATA	TCCATTAGAA	10560
55	ACAAATATTT	GAAGCTTTTA	CTTAATAGTG	ATTACCTTGA	ACTGTGCAAT	TCTAGTTTGT	10620
	AATACGTATT	TGTTTGGTTT	GTGCTTTTAT	TTTGTAAAG	TTACATTTGT	ATTATATTCA	10680
	GGAAATGAC	TTTATTATAC	TTACAGCTGT	GGTTTAAATA	CTGCTTGA	CTATTATTAT	10740
	TCITTTTACA	ACTCTTAAAG	CTTGAAGGAG	GAAGAAAAA	AAAAACAAA	CTACTAATCA	10800
	GTAGTAAATC	GAAGAGAAC	ATTTTGGCAT	TTCTTAAGAA	GAAGATGGAG	ATATTGATTA	10860
60	TATCACTTCC	TATTCAGCTG	AATAGAAAGA	ATGCCCTCAT	TGACTTGCAG	TTCTGCAGTT	10920
	TAAATTATG	AAAGAACAT	TGCTTTGCAT	TTCTGTATGA	AAGTAAAGC	ATTTTTCAGA	10980
	GAACATATG	AATTTCTCAT	ACCCAGCAGA	CAGATGGCTG	ACACTGCACA	GCCACACACC	11040
	ATTGAGTAA	GTTAAAGTGA	GAGCATAGTA	GTGAGCTCT	CCTATGAAGA	ACATTCTGGG	11100
	CTGGAGGCA	GGAATACCTC	ATGCTTGTGT	CTTTTCTCTA	CTTAAGCCCA	TTTTTGTGT	11160
65	GCTTTTCTGT	TTGTTTGTGT	TTTCACTCTT	GCATCTACGT	CTAGAGATCC	AAATGAACCTG	11220
	AAAGTTCAA	AGTTTAAACAT	ATTTAAATAT	GTTTACTTTT	AGTTGTCAAT	CTAATCGTTA	11280
	TTGATTAGAA	GCATGACTCC	TGAAGGAAG	GGAAATAAAT	CTCAATTCAT	ACTAATCTGC	11340
	AACAAACAC	TTTATCCATA	TAAATAAGTA	TATGATTAT	TTTTAAACCA	AAAAATGTAT	11400
	AAATAAAGT	TGTCCTTTAC	TGTCAATTTA	TGAGAAAGAT	CTATAATATA	TAGACTACAT	11460
70	ATATATATA	TATACAACT	AGCCAAATGT	ATGAAACCTT	GACATGTAT	AATTTGGAAT	11520
	TCACATGCTA	CCATGTAGAA	CAGGTATGAA	ATTAAAGTTAT	AATTTTCAATG	AGACATTTTC	11580
	ATCACTGTTG	ACACAGTTTC	AAGGCATTCC	ATCATGTTAT	TTTGACTTCT	TTTCTTTTTT	11640
	TTTTCTTTAA	AAATATATTT	TAACTAGAC	CAGGCCCCAC	TATAATATCA	CTTAAGAGAG	11700
	TCAGGGCAAA	GTTTGTGCT	TTATGAAGAT	GTGTTCAATG	AAGGGTGATT	GTAATGGAGT	11760
75	TCATTGGTAA	TGAAGCAAA	AGTACAGTAA	CGAAGTATTG	AAAGAAAAAT	TTTGGAGACA	11820
	TTGGAGCATA	TTATATATAG	CTTGTGGAAA	GACATAAGGC	TACAGATGGA	ATGGAACATT	11880
	OCTGTTTTCT	TGAAGAAAT	CACATACACA	TAGCTGACCT	GACTAGTACT	TCAGCTCTTC	11940
	CACAGCCTTC	TATAAAGGTT	CTTCTTCTG	CAAGAAAAAC	AAACAAAAAC	AAAAACAAA	12000
	AAAAAAAC	AAAAAAAC	AAAAAAAC	AAAAAAAC	AAAAAAAC	AAAAAAAC	12060
80	AAATACAGAA	AACAAACAC	AAAAAAGAT	TCAACCATAA	ATAGTGACTA	TTATTTTCAG	12120
	TGTGTCTTTC	ATGTGAAGAC	TATTAAGGAC	CAATATATCT	ACTGTTCATT	AGAAGAAAT	12180
	ACTTCTAAAA	CAGTAACTGA	AAATACTTAG	AGTTAAACTT	GCTGTGGATT	TTGTCTTGGC	12240
	AGTTGTCTAT	TTACATTTAT	TGTCAAAGGA	AATGTGTTTG	GCAATTAAAA	ATCTTCTCTT	12300
	AGATTIAGTG	GTGACITTA	ACCTCTTAAA	TAAATGTTAG	TATATCAGAT	TGTTGCTCTG	12360

TCACCAGATT TGGTTCTTAA ACTTT

Seq ID NO: 272 Protein sequence
Protein Accession #: NP_006465.1

```

1      11      21      31      41      51
|      |      |      |      |      |
MKVVSALLFV LGSASLWVLA EGASTGQPED DTSTTGLEGG VAMPGAKDDV VTFGTSEDRY 60
K9GLTTLVAT SVNSVTGIRI EDLPTSESTV HAQEQSPSAT ASNVATSHST EKVDGDTQTT 120
VENDGLSTVT LVGIIVGVLL AIGFIGGIIV VVMRMKSGRY SP

```

Seq ID NO: 273 DNA sequence
Nucleic Acid Accession #: CAT cluster

```

1      11      21      31      41      51
|      |      |      |      |      |
GCGGCGGCCA GCTTGCAAAG CGAAGTCTG GCCGCGCTCT TCGACTCGCT GCGCCACGTC 60
CCCCGGGGTG CCGAGCCGGC GGGGGGTGAG GTGGCTGCGC CGCGGCGCGG GCTAGGAGGT 120
CGCGGCACTG GGGGCGCGGG AGGGGAAGTG GCAGGCCCGG CGGGGGCCAC GCGGATCCCA 180
GGGCCAAGGA AGGTCCGCT CGGGGCACGC AATCTGCCTC CGTCTTCTTT CACGGAGCCG 240
TCCCGGGCAG GCGGCGCGGG GTGTGGCCCG TCGGGCCCGG ACGTGAAGCTT GGGCGACCTG 300
GAGAAGGGCG CGAGGGCCGT CGAGTTCTTT GAGCTGCTGG GCGCCGACTA CCGCGCCGGC 360
ACGAGGCGCG CAGTCTTGTCT TCGCGCCGAG CCTCTCGAGC TGTTCGCCCG CGAGGCTCTC 420
GTACTGCGGG GACCCCGCGA GCTGGAGCCC GGCCTCTTTG AGCGCCCGCC GGCAGTGGTG 480
GAAACCTTAC TGATCCCGCA GCCCTGGAGC GTCCCGGGCT GCTCCCGGAC CAAAAGAGAG 540
CCCTTGACTG CCTTCGCGCG CGGGTTGACC TTGAACGAGC CCTTGAAGCC CCTGTACCCG 600
GCGGCTGCGA ATTCTCCCG GCGGGGAGGA CGGGCCGGGC CATTGAGCTT CTTTGGCCCG 660
CTTCTTTCCA GACTGCGCTT TGC

```

Seq ID NO: 274 DNA sequence
Nucleic Acid Accession #: Bos sequence

```

1      11      21      31      41      51
|      |      |      |      |      |
CAAGAGAGGC GGGCTCCAGC TCCGGGGGTC CCGCAGTAC GGAGGCTCCG GCGGGGAACA 60
CGTGAGAGGG CTGCGCGGCA AGCAAGACTG CGGCTTCCGT GCGGCGCGCG TAGTGGGGCC 120
CCAGCAGCTC AAGAAGCTCC ACGGCTCCG CGCCCTTCTC CAGGTCCGCC AAGCTCAGT 180
CGGGCCCGCA CGGGCCACAC CCGCCGCGCG CTGCCCGGGA CGGCTCCGTG AAGAAGGAGC 240
GAGGCAGATT GCGTGCCCGC AGCGGGACCT TCGTGGCCCC TGGGATCGCC GTGGCCCCCG 300
CGGGGCGCTG CACGTCCCGT CCGCGCGCCC CAGTGCCCGC ACCTCCTAGC CCGGCGCGCG 360
GCGCAGCCAC CTCACCCCGC GCGGCTCCG CACCCCGGGG GACGTGGCGC AGCGAGTGA 420
AGAGCGGGCG CAGACTTCGG CTTTGCAAGC TGGCGGCGCG

```

Seq ID NO: 275 DNA sequence
Nucleic Acid Accession #: NM_001118.1
Coding sequence: 74..1651

```

1      11      21      31      41      51
|      |      |      |      |      |
AGCCACAGAA CACATTGGGG CTGACCTGCC GTGCTGTCA GTGGGAGGCC AGTGGTGTG 60
GCCAAGAAGT GTATGGGTG GTGTGCTGCA CGTTTCCCTG GCTGCTCACT GCGGGGCGCT 120
TCGTTGGGCG CGGGGCAGAC TCCGCAAGG ACGGCBAGCC TGCAAGTCGG CGGCCAGAG 180
ACACATTGGG GCTGACCTGC CGTGTCTGTC AGTGGGAGGC CAGTGGTGTCT GGCCAGAGAG 240
TGTGATGGCT GGTGTGCTGC ACGTTTCCCT GGTGTCTCTC CTCTGCTGTC CTATGGCCCC 300
TGCCATGGCT TCTGACTGCA TCTTCAAGAA GGAGCAAGCC ATGTGCTTGG AGAAGATCCA 360
GAGGGCCCAT GAGCTGATGG GCTTCAATGA TTCTCTTCCA GGCTGTCTGG GGATGTGGGA 420
CAACATCAGG TGTGGGAGC CCGCCATGTT GGGTGAGATG GTCTTGGTCA GCTGCCCTGA 480
GCTCTTCCGA ATCTTCAACC CAGACCAAGT CTGGGAGACC GAAACCATG GAGAGTCTGA 540
TTTTGGTGAC AGTAATCTCT TAGATCTCTC AGACATGGGA GTGGTGAGCC GGAAGTGCAC 600
GGAGGATGGC TGGTCGGAAC CCTTCCCTCA TTACTTTGAT GCCTGTGGGT TTGATGAATA 660
TGAATCTGAG ACTGGGGACC AGGATTATTA CTACCTGTCA GTGAAGGCC TCTACACGGT 720
TGGTACAGC ACATCCCTCG TCACCTCAC CACTGCCATG GTCATCTTTT GTGCTTCCG 780
GAAGCTGCAC TGACACGCA ACTTCATCCA CATGAACCTG TTTGTGTGCT TCATGCTGAG 840
GGCGATCTCC GTCTTCATCA AAGACTGGAT TCTGTATGCG GAGCAGGACA GCAACCACTG 900
CTTCATCTCC ACTGTGGAAT GTAAGGCGGT CATGGTTTTC TTCCACTACT GTGTTGTGTC 960
CACTACTTTC TGGCTGTTC TCGAGGGGCT GTACCTCTTC ACTCTGCTGG TGGAGACCTT 1020
CTTCCCTGAA AGGAGATACT TCTACTGGTA CACCATCATT GGCTGGGGGA CCCCACCTGT 1080
GTGTGTGACA GTGTGGCTA CGCTGAGACT CTACTTTGAT GACACAGGCT GCTGGGATAT 1140
GAATGACAGC ACAGCTCTGT GGTGGGTGAT CAAGGGCCCT GTGGTTGGCT CTATCATGGT 1200
TAACCTTGTG CTCTTTATTG GCATTATCGT CATCCTTGTG CAGAACTTC AGTCTCCAGA 1260
CATGGGAGGC AATGATGCA GCATCTACTT GCGACTGGCC CGGTCCACCC TGCTGCTCAT 1320
CCACTATTC GGAATCCACT ACACAGTATT TGCCTTCTCC CCAGAGAATG TCAGCAAAAG 1380
GGAAGACTC GTGTTTGAGC TGGGGCTGGG CTCCTTCCAG GGCCTTGTGG TGGCTGTCTT 1440
CTACTGTTTT CTGAATGGTG AGGTACAAGC GGAGATCAAG CGAAATGGC GAAGCTGGAA 1500
GGTGAACCGT TACTTGGCTG TGGACTTCAA GCACCGACAC CCGTCTCTGG CCAGCAGTGG 1560
GGTGAATGGG GGCACCCAGC TCTCATCTCT GAGCAAGAGC AGCTCCCAA TCCGATGTCT 1620
TGGCTCCCT GCTGACAATC TGGCCACCTG AGCCATGCTC CCTT

```

Seq ID NO: 276 Protein sequence
Protein Accession #: NP_001109.1

```

1      11      21      31      41      51
|      |      |      |      |      |
MAGVVHVSLA AHCGACFWGR GRLRRGRAAC KSAARHIGA DLPLLSVGGQ NCWPRBVMAG 60

```

	TGCGGGAGAG	CCTGGCCAAG	TGCTGCAGTT	GTTCAAGAAA	GAGAGCCCTTT	GGTGTGCTAA	420
	AGACTCTTGT	GCCCATCTTG	GAGTGGCTCC	CCAAATACCG	AGTCAAGGAA	TGGCTGCTTA	480
	GTGACGTCAAT	TTCCGGAGTT	AGTACTGGGC	TAGTGGCCAC	GCTGCAAGGG	ATGGCATATG	540
5	CCCTACTAGC	TGCAGTTCCCT	GTGGAATATG	GTCTCTACTC	TGCTTTTTC	CCTATCCTGA	600
	CATACCTTAT	CITTTGGAACA	TCAAGACATA	TCTCAGTTGG	ACCTTTTCCA	GTGGTBAATT	660
	TAATGGTGGG	ATCTGTTGTT	CTGAGCATGG	CCCCGACGA	ACACTTTCTC	GTATCCAGCA	720
	GCAATGGAAC	TGTATTAAAT	ACTACTATGA	TAGACACTGC	AGCTAGAGAT	ACAGCTAGAG	780
	TCTCTGATTG	CAGTCCCTCG	ACTCTGCTGG	TTGGAATTAT	ACAGTTGAIA	TTTGGTGGCT	840
10	TGCAGATTGG	ATTCATAGTG	AGGTACTTGG	CAGATCCITT	GGTGGTGGC	TTCAACACAG	900
	CTGCTGCCIT	CCAAGTGTCT	GTCTCACAGC	TAAAGATTGT	CCTCAATGTT	TCAACCAAAA	960
	ACTACAATGG	AGTCTCTCTT	ATTATCTATA	CGCTGGTTGA	GAITTTTCAA	AATATTGGTG	1020
	ATACCAATCT	TGCTGATTTC	ACTGCTGGAT	TGCTCACCAT	TGTCGTCTGT	ATGGCAGTTA	1080
	AGGAATTAAT	TGATCGGTTT	AGACACAAAA	TCCCACTCCC	TATTCCTATA	GAAGTAATTG	1140
15	TGACGATAAT	TGCTACTGCC	ATTTCATATG	GAGCCAACTT	GGAAAAAAT	TACAATGCTG	1200
	GCATTTGTAA	ATCCATCCCA	AGGGGGTTTT	TGCTCTCTGA	ACTTCCACCT	GTGAGCTTGT	1260
	TCTCGGAGAT	GCTGGCTGCA	TCAITTTTCCA	TGCTGTGGTT	GGCTTATGCT	ATTGCAAGTG	1320
	CAGTAGGAAA	AGTATATGCC	ACCAAGTATG	ATTACACCAT	CGATGGGAAC	CAGGAATTCA	1380
	TTGCGCTTGG	GATCAGCAAC	ATCTTCTCAG	GAITTTCTCT	TTGTTTGTG	GCCACCACTG	1440
	CTCTTTCCCG	CACGGCCGTC	CAGGAGAGCA	CTGGAGGAAA	GACACAGGTT	GCTGGCATCA	1500
20	TCTCTGCTGC	GATTTGTGATG	ATCGCCATTC	TTGCCCTGGG	GAAGCTTCTG	GAACCCCTGC	1560
	AGAACTCGGT	CTTGGCAGCT	GTGTAATTG	CCAACTGTAA	AGGGATGTTT	ATGCAGCTGT	1620
	GTGACATTC	TGCTGTTGG	AGACAGAAIA	AGATTGATGC	TGTTATCTGG	GTGTTTACGT	1680
	GTATAGTGTG	CATCATTTCTG	GGGCTGGATC	TGCGTTTACT	AGCTGGCCTT	ATATTGGGAC	1740
25	TGTTGACTGT	GTCTCTGAGA	GTTCAGTTTC	CTTCTTGGAA	TGCGCTTGGG	AGCATCCCTA	1800
	GCACAGATAT	CTACAAAAGT	ACCAAGAAAT	ACAAAACAT	TGAAGAACCT	CAAGGAGTGA	1860
	AGATTCTTAG	ATTTTCCAGT	CCTATTTTCT	ATGGCAATGT	CGATGGTTT	AAAAATGTGA	1920
	TCAGATCCAC	CGTTGGATTG	GATGCCATTA	GAGTATATAA	TAAGAGGCTG	AAAGCGCTGA	1980
	GGAAAAATCA	GAACATAATA	AAAAGTGGAC	AATTAAGAGC	AACAAAGAAAT	BGCATCATAA	2040
30	GTGATGCTGT	TTCAACAAT	AATGCTTTTG	AGCCTGATGA	GGATATTGAA	GNTCTGGAGG	2100
	AACCTGATAT	CCCAACCAAG	GAATAGAGA	TTCAAGTGGG	TTGGAACCTC	GAGCTTCCAG	2160
	TCAAAGTGAA	CGTTCCTAAA	GTGCCAATCC	ATAGCCTTGT	GCTTGACTGT	GGAGCTATAT	2220
	CTTTCCTGSA	CGTTCTTGGA	GTGAGATCAC	TGCGGTTGAT	TGTCAAAGAA	TTCCAAAGAA	2280
	TTGATGTGAA	TGTGATTTT	GCATCCTTC	AAGATTATGT	GATAGAAAG	CTGGAGCAAT	2340
35	GCGGTCTTGT	TGAGACCAAC	ATTAGAAAGG	ACACATTCIT	TTTBACTGGT	CTGATGCTA	2400
	TACTCTATCT	ACAGAACCAA	GTGAAATCTC	AAGAGGGTCA	AGGTTCCATT	TTAGAAACGA	2460
	TCACTCTCAT	TCAGGATTGT	AAAGATACCC	TTGAATTAAAT	AGAAACAGAG	CTGACGGGAG	2520
	AAGAAGTTGA	TGTCAGGAT	GAGGCTATGC	GTACACTTGC	ATCCGAAAG	TGGGTTCCGG	2580
	AGGTCTCTAT	GAGCAAGGAA	TACAAGACAA	AACCTTCTCA	ATGCATTGAC	TATTTCTTCA	2640
40	GACTCAAAAC	ACTCATCTCT	TTTTCTATTA	AGCCATTGAA	AGAGAAGCAC	TAAGACTGCT	2700
	TCTAGGCTTT	ATTTATAAAA	TAAACACCTT	ATCCCTAACA	TGGGCAAAAT	GGCTAGAAAT	2760
	ATTACAGAGA	TTTGGCAGCG	TCCAGGGTAA	GCTGGTGTTA	TAATACGCTG	CTGATCTACA	2820
	TCACAGATT	GCTAATAATG	TTCACTGGG	CCCTGGCATA	TCTCTGTCA	GTTAGAGTGA	2880
	GTGCTGACCC	AACAGCCTCT	GTGGTCAAGC	GAGTCACGAA	TGATTAATCA	TAAAGAAATA	2940
45	TCAGTTTCTG	TGAGACCTGG	ATATCCATGA	GCTGCACCTG	TCACCATGTA	AGGTCACATT	3000
	TAGTAATGTC	TGAATAAAAA	TGATTAAGTC	ATTTATCAAT	AAAAGCCITT	GAATAACTTT	3060
	TGGTAATAAA	ATTGGAGTTT	TAAAAATGCA	AATTTGCTTA	GTATCTAATA	ATGAAGTGT	3120
	ATTACATATA	GCCGGAAATG	AGGATCTCTT	TGATCCCTGGA	AATGGTTTAC	CTAAAGAGCTA	3180
	CAGAACCCAG	CCAATATATT	TTGAAATATT	GATGCAAGCA	AATGAATAAA	TAAAGAGATT	3240
50	TTCAATGTTT	ATAAAGATCT	TTTTTGATAT	GATAATAATC	ATGATCACAA	CTGAGATCAA	3300
	AAAAATATAT	GACAGATTAT	TTTGTTTAAA	AATGCAGTTT	TAATTATCTT	AGTCTATAGA	3360
	AATGATCAIT	GCATGAGGCT	ATGATATAGT	ATGATCTGTG	TAAATCTGTA	CATAAAAAAC	3420
	GTGCTATTCT	GAGTGAAAT	TTTTTTGATG	TGCTTACATA	ACCATGGTGA	TTAAATGAG	3480
	TTTATATTTT	TTCTCAAAA	TTTTAGCAGT	GTGTAAAGTA	AGTAATCTTT	AACCTGAATC	3540
55	TGACCACTTA	AAAAAAATC	ACTACCTATA	GTAGTCTGTG	TTTAAAGTGA	TTTAAAGTGA	3600
	ATTTTAAAG	ACAAAGCAAT	CTAAATGAAC	TCAATATAAA	AACATTCATT	TGGAATGTAC	3660
	ATCTGAAATA	ATACAGGTTT	TTTTGACCAA	AGGTTTTTAT	ATCTTTTCTT	TTTATTATT	3720
	TTTTTCTTAA	GTGCAACAA	TTTTCTAGAT	ATTATATACA	ACACAGGCTT	TGATCTTGGG	3780
	GACTTTTCCC	ATATATTCCA	CACCTGGAGT	AATGAAGTTG	TACTTCAATT	CTAGAGAAAA	3840
60	GTATACCTCA	GGTCCCAAT	TGAGAAATGC	TTGCTTGATT	GAAGACGACA	TCATCCCTTG	3900
	GTATACCTCA	GGGATTGGTT	TCAGGACCCC	TGCATTATCC	AAAATTTGTG	CACACTCAG	3960
	TCCTGCGATC	ACCCCTGGCT	AAAGATAGAA	TGGCTTCTCT	GTTTTCTCTC	TGAATACAA	4020
	CCAGAAACAA	TGTGTCTATT	TCTGAAAGAA	TAGGATTAAT	GATCATACAA	ATGGGTTAAT	4080
	CCGTAAATCT	GTTGTAAAT	CTGGTTACAG	CATAACTAGG	ATTATAATGC	TGCCCTCATT	4140
65	TCACAGCAT	ATTGCTTAT	ATTGACAA	AATCATCTCG	CTAAAGAGTG	AATGTAGGCC	4200
	AGGCGCGGTG	GCTCATGCTT	GTAAATCCAG	CACTTTGGGA	GGCCGAGGCG	GGTGGATCAC	4260
	GAGGTCAGGA	GATCGAGACC	ATCCTGGCTA	ACATGGTAAA	ACCCGCTCTC	TACTAAAAAT	4320
	AGAAAAAAG	AAATTAGCCT	AGCGTGCTGG	CTGGCGGGCG	CCTGTAGTCC	CAGCTATTGG	4380
	GGGGCTAAG	CGAGGAGAAAT	GGCGTGAACC	CGGGAGGCGG	AGCTTGCACT	GAGCCGAGGT	4440
70	CGTGCCACTG	CACCTCAGCC	TGGGCGACAG	AGCAAGACTC	CGCTCAAAA	AAAAAATAAA	4500
	AAAAAATAAA	AGAGTGAATG	TAATAGTCTT	GCAGAAATG	AATGAATACC	TTTGTTCAT	4560
	AAAGGAATTA	TGCACTGCTC	ACTTTTGTGA	AGGAAATGCC	AAAGTATAGT	TTTACACAAA	4620
	GGCTAGAGTT	TGTAAATTTT	GGGTTCAATT	GTGATGACAT	AGGTCAGCAA	ACTGCGGGAA	4680
	TACTGTCTCT	TCTATGATTT	TTTGAATAG	TAAGCATAAT	TTTAGTTTGG	TATTATCAAT	4740
75	GAATAATTCA	CTTGAAATTA	AAAGTGCCTT	TTGTATATTT	TTTAACTTAT	AGGATAAGAT	4800
	TCCAGTATG	TATATGAGTT	TTAACAAAT	AAAAATCAA	ATCATGTACA	TTTGAATAA	4860
	TTTGACACAA	TTTAAAAATA	AATGTAAAGT	TGTTCTTTAA	ACTACTCGGA	TGTGTCTTT	4920
80	CTGAACAAAA						

A212 Protein sequence:

Gene name:	Solute carrier family 26, member 4
Unigene number:	Hs.159275
Probeset Accession #:	AF030880
Protein Accession #:	O43511
Signal sequence:	none found

Transmembrane domains: 81-103, 109-131, 136-158, 185-207, 221-243, 245-267, 270-291, 295-317,
347-369, 386-408, 420-442, 448-470, 486-508
Cellular Localization: plasma membrane

5

10

15

20

```

1      11      21      31      41      51
|      |      |      |      |      |
MAAPGGRSEF POLPEYSCSY MVRPVSYSKL AFQQQHERRL QERKTLRESL AKCCSCSRKR 60
APGVLTSLVP ILEWLKPYRV KENLLSDVIS GVSTGLVATL QGMAYALLAA VPGYGLYSA 120
FPPIITYFIP GTSRHISVGP FVVSLMVGS VVLEMAPDEH PLVSSSNGTV LNTTMDTAA 180
RDTARVLIAS ALTLVLGIIQ LIFGGLQIGF IVRYLADPLV GGFTTAAAFQ VLVSQKLKIVL 240
NVSTKXNGV LSIIYTLVEI FQNIQDYNLA DFTAGLLTIV VCMAYKELND RFRHKIPVPI 300
PIEVIVTIIA TAISYGANLE RNYNAGIVKS IPRGFIPPEL PFVSLFSEML AASFSLAVVA 360
YALAVSVGVK YATKYDYTID GNQEFIAFGI SNIFSGFPSC FVATTALSRT AVQESTGGKT 420
QVAGIISAAI VMIALLAGLK LLEPLQKSVL AAVVIANLKG MFMQLCDIPR LWRQNKIDAV 480
INVFTCIVSI ILGDLGLLLA GLIFGLLTIV LRVQFESWNG LGSIPSTDIY KSTKNYKNIE 540
EPQGVKILRF SSPIFYGNVD GPKKCIKSTV GFDAIRVYNK RLKALRKIQK LIKSGQLRAT 600
KNGIISDAVS TNNAFEDDED IEKLEELDIP TKEIEIQVDW NSRLPVKVMV PKVPIHSLVL 660
DCGALSFLDV VGVRSRLRVIV KEFQRIDNVN YFASLQDYVI EKLEQCGPFD DNIRKDTFFL 720
TVHDAIYLQ NQVKSQBGQG SILETITLIQ DCKDTLELIE TBLTEELDVI QDEAMRTLAS 780
QDEAMRTLAS

```

25

30

A213 DNA SEQUENCE:

Gene name: ESTs, Highly similar to calcium-activated potassium channel rSK2 [R.norvegicus]
Unigene number: Hs.98280
Probeset Accession #: AA418000
Nucleic Acid Accession #: NM_021614
Coding sequence: 458-2197 (underlined sequences correspond to start and stop codons)

35

40

45

50

55

60

65

70

75

```

1      11      21      31      41      51
|      |      |      |      |      |
CGGCGGCAGC AGCCCATGCC TCCGGTGCAA CAGCTGCGCC TCCTCCGGTG CCCC GGCGGC 60
GGGGGCGGGA GATAACCTGT CCTGCTGCTC CCGCACTCC TCGCCCGGCG GCGCCTCCG 120
GACCGGCACC TCCTCGCCGC TGTCGGGCTC GTCTCTCTGC TGCTGCTGCT GCTCGTCCG 180
CCGGGCGAGC CAGCTCAATG TGAGCGAGCT GACGCGGTCC AGCCATGCCA GTGGCGTCC 240
GCAGCAGTAC GCGCAGCAGT CCGCGCAGCA GTCGGGTCC GCCTCCCGT ACCACCAATG 300
CCACAGCGTG CAGCCCGCGG CCGACCCACC GGGCAGCGTC GGCAGTCTCG GCTCCGCGCC 360
CCGCTCTCG CACCCACCAC ACCACCCGCA CCGCGCGCAC CACCCAGACC ACCAGCCCA 420
GGGCGCGCGC GAGAGCAACC CTTTCAACGA AATAGCATG AGCAGCTGCA GGTACAACGG 480
GGCGCTCATG CCGCGCTCA GCAACTTGAG CGGCTCCGCG CGGAACCTCC ACGAGATGGA 540
CTCAGAGGCG CAGCCCTGCG AGCCCTCCGC GTCTGTGCGA GGAGGTGGGG GCGCTCTCTC 600
CCGCTCTGCA GAGCGTGGCG CCGCGCGCGC TGTITGCTCC TCAGCCCGCG AGATCGTGGT 660
GTCTAAGCCC GAGCAACA CACTCCACAA CTTGGCGCTC TATGGAAACG GTGGCGGAGG 720
CAGCACTGGA GGAGGCGGGG GCGGTGGAGG GAGCGGGCAC GGCAGCAGCA GTGGCACCAA 780
GTCCAGCAAA AAGAAAAACC AGAACATCGG CTACAGCTG GGCACCGCGC GCGCCTGTT 840
CGAAAAGCGC AAGCGGCTCA GCGACTACGC GCTCATCTTC GGCATGTTCC GCATGTRGGT 900
CATGGTCATC GAGACCGAGC TGTCGTGGGG CGCCTACGAC AAGGCGTCCG TGTATTCCTT 960
AGCTCTGAAA TGCTTATCA GTCTCTCCAC GATCATCTCG CTCGGTCTGA TCATCGTGA 1020
CCACGCGCAG GAAATACAGT TGTCATGCT GGCATGGA GCAGATGACT GGAAGATAGC 1080
CATGACTTAT GAGCGTATTT TCTTCATCTG CTTGGAATA CTGGTGTGTG CTATTCATCC 1140
CATACCTGGG AATATACAT TCACATGGAC GGGCGCGCTT GCCTTCTCT ATGCCCATC 1200
CACAAACCAC GCTGATGTGG ATATTATTTT ATCTATACCA ATGTTCTTAA GACTCTATCT 1260
GATTCGCAGA GTCATGCTTT TACATAGCAA ACTTTTCACT GATGCCCTCT CTAGAAGCAT 1320
TGGAGCACTT AATAAGATAA ACTTCAATAC AGTTTGTGTT ATGAAGACTT TAATGACTAT 1380
ATGCCAGAGG ACCTGACTCT TGGTTTITAG TATCTCATT TGGATAATTG CCGCATGGAC 1440
TGTCCGAGCT TGTGAAGGT ACCATGATCA ACAGGATGTT ACTAGCAACT TCCTTGGAGC 1500
GATGTGGTTG ATATCAATAA CTTTCTCTC CATGGTTAT GGTGACATGG TACCTAACAC 1560
ATACTGTGGA AAGGAGTCT GCTTACTTAC TGGAAATATG GGTGCTGGTT GCACAGGCT 1620
GGTGTAGCT GTAGTGCAA GGAAGCTAGA ACTTACCAA GCAGAAAAAC ACGTGCACAA 1680
TTTCATGATG GATACTCAGC TGACTAAAAG AGTAAAAAAT GCAGCTGCCA ATGTACTCAG 1740
GGAAACATGG CTAATTACAA AAAATACAAA GCTAGTGAAA AAGATAGATC ATGCAAAAGT 1800
AAGAAAACAT CAACGAAAT TCCTGCAAGC TATTCATCAA TTAAGAAAGT TAAAAATGGA 1860
GCAGAGGAAA CTGAATGACC AAGCAAACAC TTTGGTGGAC TTGGCAAAGA CCCAGAACAT 1920
CATGTATGAT ATGATTTCG ACTTAAACGA AAGGAGTGAA GACTTCGAGA AGAGGATTGT 1980
TACCCCTGGA ACAAACCTAG AGACTTGTAT TGGTAGCATC CACGCCCTCC CTGGGCTCAT 2040
AAGCCAGACC ATCAGGCAGC AGCAGAGAGA TTTCAITGAG GCTCAGATGG AGAGCTACGA 2100
CAGACGCTC ACTTACATG CTGAGCGGTC CCGTCTCTCG TCCAGGAGGC GCGGCTCTCT 2160
TTCCACAGCA CCACCAACTT CATCAGAGAG TAGCTAGAAG AGAATAAGTT AACCAAAA 2220
TAAGACTTTT TGCCATCATA TGGTCAATAT TTTAGCTTTT ATGTAAAGC CCTATGGTT 2280
CTAATCAGCG TTATCCGGGT TCGATGTCA GAATCTGGGG AACCTGAACA CTAAGTTTIA 2340
GGCCAAATG AGTGAAACT CTTTTTTTTT CTTTCAGATG CACAGGGAAT GCACCTATTA 2400
TGTCTATATA GATGTTCCT CCGTAAATTT CACTACATTT TTATTCATGC ACTTCAACA 2460
AACTTTACTA CTACATTATA TGATATATAA TAAAAAAGT TAATTTCCGA

```

80

A214 Protein sequence:

Gene name: ESTs, Highly similar to calcium-activated potassium channel rSK2 [R.norvegicus]
Unigene number: Hs.98280
Probeset Accession #: AA418000
Protein Accession #: NP_067627
Signal sequence: none found
Transmembrane domains: 135-157, 168-190, 208-230, 254-276, 306-328, 342-364, 373-394

Calmodulin binding domain: 412-488
Cellular Localization: plasma membrane

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
MSSCRYNGGV MRPLSNLSAS RRNLHMDSE AQPLQPPASV GGGGGASSPS AAAAAAASVS 60
SSAPEIVVSK PERHNSNNLA LYGTGGGGST GGGGGGGGGG HGGSSGTSKS KKKNNIGYK 120
LGERRALFEK RRLSDYALI PQMPGIVVMV IETELSWGAY DKASLYSLAL KCLISLSTII 180
10    LLGLIIVYHA REIQLFMVDN GADDWRLAMI YERIFFICLE ILVCAIHPIF GNYTFTWTAR 240
LAFSYAPSTT TADVDIILSI PMFLRLYLIA RVMLLHSLKF TDASSRSIGA LNKINPNTRF 300
VMKTLMTICP GTVLLVFSIS LWI IAAWTVR ACERYHDQQD VTSNPLGAMW LISITFLSIG 360
YGDMPVNTYC GKGVCLLTGI MGAGCTALVV AVVARKLELT KAEKHVHNFN MDTQLTKRVK 420
NAAANVLRET WLIYKNTKLV KKIDKAKVRK HQKFLQAIH QLRSVKMEQR KINDQANTLV 480
15    DLAKTQNIIM DMISDLNERS EDFEKRICTL ETKLETLIGS IHALPGLISQ TIRQQQRDFI 540
EQQMESYDKH VTYNAERERS SSRRRSSST APPTSSESS

```

A215 DNA SEQUENCE:

Gene name: CGI-86 protein
Unigene number: Hs.109201
Probeset Accession #: AN161450
Nucleic Acid Accession #: NM_016029
Coding sequence: 228-1097 (underlined sequences correspond to start and stop codons)

```

25      1      11      21      31      41      51
      |      |      |      |      |      |
CTGCGATCCC GCAGGGCAGC GACGCGACTC TGGTGCGGGC CGTCTTCCTC CCCCGAGACT 60
GGGCGTGGGC GGCSCAAATG AACGCGGAGC TGCTGCTGTG CTGCTGTGTG CTGTGCGCGC 120
30    TGCTCCTGCT CTGCTGTGAG CTGCTGCGCT TCCTGAGGGC TGACGCGGAC CTGACGCTAC 180
TATGGGCCGA GTGGCAGGGA CGACGCCGAG AATGGGAGCT GACTGATATG GTGGTGTGGG 240
TGACTGGAGC CTCGAGTGGA ATTGGTGAGG AGCTGGCTTA CCAGTTGTCT AAACCTAGGAG 300
TTTCTCTTGT GCTGTGAGCC AGAAGAGTGC ATGAGCTGGA AAGGGTGAAA AGAAGATGCC 360
TAGAGAATGG CAATTPTAAA GAAAAAGATA TACTTGTTTT GCCCCTTGAC CTGACCGACA 420
35    CTGGTTCCCA TGAAGCGGCT ACCAAAGCTG TTCTCCAGGA GTTGTGTAGA ATCGACATTC 480
TGGTCAACAA TGCTGGAATG TCCGAGCGTT CTCTGTGCAT GGATACGAGC TTGGATGTCT 540
ACAGAAAGCT AATAGAGCTT AACTACTTAG GGACGGTGTG CTTGACAAAA TGTGTTCTGC 600
CTCACATGAT CGAGAGGAAG CAAGGAAAGA TTGTTACTGT GAATAGCATC CTGGGTATCA 660
TATCTGTACC TCTTTCCATT GGAATACTGT CTAGCAAGCA TGCTCTCCGG GGTTTTTTTA 720
40    ATGGCCTTCG AACAGAATCT GCCACATACC CAGGTATAAT AGTTTCTAAC ATTGCCCCAG 780
GACCTGTGCA ATCAAAATAT GTGGAGAATT CCTAGCTGCG AGAAGTCACA AAGACTATAG 840
GCATAATGCG AGACCAATCC CACAGATGA CAACCACTCG TTGTGTGCGG CTGATGTTAA 900
TCAGCATGGC CAATGATTGG AAAGAAGITT GGATCTCAGA ACAACCTTTC TTGTTAGTAA 960
45    CATATTGTG CAAATACATG CCAACCTGGG CCTGGTGGAT AACCAACAGG ATGGGGAAGA 1020
AAAGGATTGA GAACCTTAAG AGTGGTGTGG ATGCAGATC TTCTTATTTT AAAATCTTAA 1080
AGACAAACAA TGACTGAAAA GAGCACCTGT ACTTTTCAAG CCACTGGAGG GAGAAATGGA 1140
AAACATGAAA ACAGCATCTT TCTTATGCTT CTGAATAATC AAAGACTAAT TTGTGATTTC 1200
50    ACTTTTAAAT AGATATGACT TTGCTTCCAA CATGGAATGA AATAAAAAAT AAATAATAAA 1260
AGATTGCCAT GAATCTTGCA AA

```

A216 Protein sequence:

Gene name: CGI-86 protein
Unigene number: Hs.109201
Probeset Accession #: AN161450
Protein Accession #: NP_057113
Signal sequence: 1-26
Transmembrane domains: 183-206, 221-243
Cellular Localization: plasma membrane

```

60      1      11      21      31      41      51
      |      |      |      |      |      |
MNWELLLMLL VLCAILLILLV QLLRFLRADG DLTLLWAEMQ GRRPEWELTD MUVWVTGASS 60
65    GIGKELAYQL SKLGVSLLVLS ARRVHELERV KRRCLENGSL KEKDILVLPD DLTDTGSHEA 120
ATKAVLQEPG RIDILVNNGG MSQRSLCMDT SLDVYRKLIE LNYLGTVSLT KCVLPHEMER 180
KQGIIVTVNS ILGIISVPLS IGYCASEHAL RGFFNGLRTE LATYPGLIIV NICPGPVQSN 240
IVENSLAGEV TKTIGNNGDQ SHEMTTSCRV RIMLISMAND LKEVWISBEP FLVVTYLVQY 300
70    MPTWAWWITN KMKKKRIENT KSGVDADSSY FKIPKTKED

```

A217 DNA SEQUENCE:

Gene name: Homo sapiens mRNA; cDNA DKFpZp56401763 (from clone DKFpZp56401763)
Unigene number: Hs.27373
Probeset Accession #: F13036
Nucleic Acid Accession #: AC012478
Coding sequence: 1-1896 (underlined sequences correspond to start and stop codons)

```

80      1      11      21      31      41      51
      |      |      |      |      |      |
ATGCGCGCGG TGCCGCTGCG CGCCCCGCTC CTGCGCGCTG TCGTCTGCGC GCTCCTGGCC 60
GCTCCGCGCG CCGCGCGCAG CAGAGCGGAG TCCGTCTCCG CGCGCGTGGC CGAACCOCAG 120
CGCGAGTCGC GGCACACGCC CGGCCCGGGG CCGCGGACAA CCACCGCGTT TGGGTCTGGG 180
GCGCGCGGCG GCAGCGGCAG CTCACGCTCC AACAGCAGTG GCGACGCTTT GGTGACCCGC 240

```

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

```

ATTTCATCC TCCTCCGCGA CCTACCCACC CTCAAGGCAG CGGTGATCGT GGCCTTCGCC 300
TTTACCACCC TCCTCATCGC CTGCTGCTGT CTGCGGCTCT TCAGGTTCGGG AAGAGAGTTA 360
AAGAGACACG GCAAGTATGA TATCATCACC ACTCCAGCAG AGCGAGTGGG AATGGCGCCA 420
CTAAATGAAG AGGATGATGA AGATGAGGAC TCACAGTAT TCACATCAA ATACAGAGTG 480
TCCTTGCGCG CTGCACTGAG ACCTCAGCTG CCAAGGTGCC AGACGCTACT GACAGTTCCT 540
GTGCCCCCAC CCTTCATCCT CGACATTGAC CTTCCAGCAA GATGCACTGG AAGGCTTGAT 600
GGTGGATCA GACCTGGTAA AACCTGTTTC CCAGCCTGGT GGCATCCTGT GGAAGATTGG 660
TCAGCTGCAA CTCTGGGTGT GAAGGACTGG ACCTGGGAAG CCTCTGCGT CGGAGGTGTT 720
GAAACCAAAA CGAACGTTAT GTATAAAACC CCAGCTCCAT CGTGCCTGTC AGGCATCTGC 780
TCAGACTGTC ACTGGCAAGC TCGTTTCACG GTCACCAAA TGGAGTTGCT TCTGCCACCC 840
TTTGGGCATC CCTTTAAAGT GCCCCCTACT TCTACTCCCC ATGCTTTTCG ACAACTGCAG 900
CTGAATGCTA TGGAAAAGCT GGATTCTCTT GCCTTACGCA GAAACACCGG GGCTCCATCT 960
GCCAGGTGCT TGCCACTGGT CCTGGCAGAA ATGGCGGCTG CTGAAAGTGA CCTTCCAAAT 1020
CCTTGCTGGC ACTTCAGGCG CACAGSCTCT CCAATAAAAA CCTTTACAC ACNACCATG 1080
AGTACCTTGG GCTTGGATGT TTTCTGTGGT GCCGCCACG GGGGCCACCT TTTGGAAGAC 1140
AGAGCAGTGA CTAAGTTCCT CCAGGCTAGC TCTTTCTCCA AACAGCTGGG CTGGAAGCCA 1200
GCCTAGAGA GTGGGTGTTT CCATCATCTC AGGCTTCTCA GAGAGTGTCC TCCGCTGAGC 1260
ACCATCCTG TCAGTTTGGC TCGTTGAGT GCCCGGGGAC AAGCCAGCCT GACCGGGAGG 1320
AGGCTGTTC GCGCTCCGGG GCAGTCTCTG CATGGCGGAG GGTACAGCGG TACCGCACT 1380
TGCCCTTTGG TTTTGAAGAT TCTGTGAGG CGCCATCTCT ACCTTGACCT CTTCTACAAA 1440
ATCTGTCTCC CTGTCTGTGC CGTGGAAACAC CTACGGGAAG CCAAGAGAAG CTCAGTGACT 1500
GTCTTGCCTA GTGGGTGACA GAGCCACAAA AAGGCAGCTG CTGCCACCGG GGAGCCTGTC 1560
AAGCGAGGCG CCAGTGGGCA ATTGACCAAG CACACATGCC CTGGCTGGGG GATCACACAT 1620
GCGACCTGCG AGACATTCCT AGATACCCAA GGCAGGAAG GCCACGCTGA GGTATGCTCT 1680
CACCTGGAG GAGACTTGA TGGGTGGCA AATTTCATT TGGAGGAAGA GGTTCCTCAG 1740
GATGGCAGAT GCCAGAGAT GGTCTGTGAT TCTGAGGAAG GGCCACTAG TTTGACAGGA 1800
TGTGAGAGGC TCACAGTTC CCATCACTTC TCAGGCCATT CCAAGTCTTG GTCTTCTCTT 1860
TCCCCCGGAC AGCCCCGTGT TCTGTCCAGG CCCTGA
  
```

A218 Protein sequence:

Gene name: Homo sapiens mRNA; cDNA DKFZp56401763 (from clone DKFZp56401763)
 Unigene number: Hs.27373
 Probeset Accession #: FL3036
 Protein Accession #: P0ENESH predicted
 Signal sequence: 1-27
 Transmembrane domains: 94-115, 448-469
 Cellular Localization: not determined

1 11 21 31 41 51
 MRVPLPAPL LPLLLALLA APAARASRAE SVSAFWPEPE RSRPPPPHFG PNTTRFGSG 60
 AAGSGSSSS NSSGDALVTR ISILLRDLPT LKAAVIVAF FTLILLACL LRVFRSGKRL 120
 KKTRKYDIIT TPAERVMAP LNEEDDEDED STVFDIKYRV SLPAALERQL PGQQLLTVP 180
 VPPFFILDI LPARCSGRPD GDIRPGKTCF PAWHHPVESW SAATGWGKDW TNKPCVGGV 240
 ETKINVMYKT PAPSCVSGIC SDCHWQARFE VTMELLLEPP FGHFFKVPPT STPHGFRQLQ 300
 INLMKLDSS ALRRNTRAPS ARCLFLVLAE MAAABSDLEN PWNHFBATGS PIKILYTOTM 360
 STLGLDVFCED AGQRTGFCED RAVTKVLQGS SFSKQLRWKP ALESGFHHHL RLARECFPLS 420
 THPVRLARSD ARGQASLTGR RVFRRPRQSL HGGGSAGTAT CILVLKILLR RHPFLDLFYK 480
 ICLPCCAVEH LREAKRSSVT VLASFQSPQ KAAAABGEPP KRBPFGQLTR HTCPGWGITH 540
 ANLQTIPTDQ GQSGPRDVT HPGDLDGVA NFYLEEGEFQ DGRQKQMLVM SEEGPPSLTG 600
 CERLTGSHHF SSESKSWSFL SPRQLFLSR P

A219 DNA SEQUENCE

Gene name: selectin E (endothelial adhesion molecule 1)
 Unigene number: Hs.89546
 Probeset Accession #: M24736
 Nucleic Acid Accession #: NM 000450
 Coding sequence: 1-1833 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51
 65
 70
 75
 80

```

ATGATTGCTT CACAGTTTCT CTCAGTCTCT ACTTTGTCG TCTCATTA AGAGAGTGGG 60
GCCTGGTCTT ACAACACCTC CACGGAAGCT ATGACTTATG ATGAGGCCAG TGCTTATTGT 120
CAGCAAGGT ACACACACCT GGTGCAATT CAARAACAG AAGAGATTGA GTACCTAATC 180
TCCATATTGA GCTATTACAC AAGTTATTAC TGGATTGGAA TCAGAAAAGT CACCAATGTG 240
TGGGTCTGGG TAGGAACCCA GAAACCTCTG ACAGAAAGAG CCAAGAACTG GGCTCCAGGT 300
GAACCAACA ATAGGCAAAA AGATGAGGAC TCGTGGAGA TCTACATCAA GAGAGAAAAA 360
GATGTGGGCA TGTGCAATGA TGAAGGTGTC AGCAGAAGA AGCTTGCCCT ATGCTACACA 420
GCTGCTGTA CCAATACATC CTGCACTGGC CACGGTGAAT GTGTAGAGAC CATCAATAAT 480
TACACTTGA AGTGTACACC TGGCTTCAGT GGACTCAAGT GTGAGCAAT TGTGAATCTG 540
ACAGCCCTGG AATCCCTGGA GCATGGAAGC CTGGTTTGA GTCAACCACT GGGAACTTTC 600
AGCTACAATT CTCTCTGCTC TATCAGCTGT GATAGGGGTT ACCTGCCAAG CAGCATGGAG 660
ACCATGCACT GTATGTCTTC TGGAGAATGG AGTGCTCCTA TTCCAGCCTG CAATGTGGTT 720
GAGTGTGATG CTGTGACAAA TCCAGCCAAT GGGTTCGTGG AATGTTTCCA AAACCTTGA 780
AGCTTCCCAT GGAACACAAC CTGTACATTT GACTGTGAAG AAGGATTGTA ACTAATGGGA 840
GCCAGAGGCC TTCACTGTAC CTCACTGGGG AATTGGGACA ACAGAAAGCC AACGTGTAAA 900
GCTGAGACT CAGGGGCCGT CGGCAGCCT CAGAATGGCT CTGTGAGGTG CAGCCATTCC 960
CTGCTGGAG AGTTCACTT CAAATCATCC TGCAACTCA CCGTGTGAGA AGGCTTCATG 1020
TTGACGGGAC CAGCCAGGT TGAATGCACC ACTCAAGGGC AGTGGACACA GCAATCCCCA 1080
GTTTGTGAAG CACAGCCTTG TCCAACCCCG AGCGAGGCTA CATGAATTGT 1140
CTTCTAGATG CTTCTGGCAG TTTCCGTTAT GGTCCAGCT GTGAGTCTC CTGTGAGCAG 1200
GGTTTGTGT TGAAGGATC CAAAGGCTC CAATGTGGCC CCACAGGGA GTGGACAAC 1260
  
```


5
10
GAGAAGCCCA CATGTGAAGC TGTGAGATGC GATGCTGTCC ACCAGCCCCC GAAGGGTTTG 1320
GTGAGGTGTG CTCATTCCCC TATTGGAGAA TTCACCTACA AGTCCTCTTG TGCCCTTCAGC 1380
TGTGAGGAGG GATTGTGAATF ATATGGATCA ACTCAACTTG AGTGACATC TCAGGGACAA 1440
TGACAGGAAG AGGTTCCCTC CTGCCAAGTG GTAAAAATGTT CAAGCCTGGC AGTTCCGGGA 1500
AAGATCAACA TGAGCTGCAG TGGGGAGGCC GTGTTTGCCA CTGTGTGCAA GTTCGCTGT 1560
CCTGAAGGAT GGAGGCTCAA TGGCTCTGCA GCTCGGACAT GTGGAGCCAC AGGACACTGG 1620
TCTGGCCTGC TACCTACCTG TGAAGCTCCC ACTGAGTCCA ACATTCCTCT GGTAGCTGGA 1680
CTTCTGCTG CTGACTCTC CTCTCTGACA TTAGCAOCAT TTCTCTCTG GCTTCGGAAA 1740
TGCTTACGGA AAGCAAAGAA ATTGTTCCT GCCAGCAGCT GCCAAGCCT TGAATCAGAC 1800
GGAAGCTACC AAAAGCCTTC TTACATCCTT TAA

A220 Protein sequence:

15
20
Gene name: Selectin E (endothelial adhesion molecule 1)
Unigene number: Hs.89546
Probeset Accession #: M24736
Protein Accession #: NP_000441
Signal sequence: 1-22
Transmembrane domains: 555-573
C-lectin domain: 23-139
Cellular Localization: plasma membrane

25
30
35
1 11 21 31 41 51
MIASQFLSAL TVLLIKESG ANSYNTSTEA MYDEASAYC QQRVTHLVAI QNKEIEIYLN 60
SILSYSPSYW WIGIRKVNNV WVVGTQKPL TBAKNWAPG EPNNRQKDED CVEIYIKREK 120
DVGMWNDERC SKKLALCYT AACTINTSCSG HGECEVETINN YTCCKDPPGFS GLKCEQIVNC 180
TALESFENGSL VCSHPLGNF SYNSSCSISC DRGYLPSEME TMQCMSSGSM SAPIPACNVV 240
ECDAVTPANW GFVECFQNEG SFPWNTTCTF DCEBGFELMG AQSLQCTSSG NWDNEKPTCK 300
AVTCRAVROP QNGSVRCSSHS PAGEFTFKSS CNFTCEBGM LQGPQVECT TQGGWTQIIP 360
VCEAFQCTAL SNFERGYMNC LPSASGSEFRY GSSCEFSCEQ GFVLKGSKRL QCGPTGWDN 420
EKPTCEAVRC DAVHQPFGKL VRCASPTIGE FTYKSSCAPS CEBGFELYGS TQLECTSQGQ 480
WTEVPSCQV VRCSSLAIVEG KINMSCSGEP VFGTVCKFAC FEGWTINGSA ARTCGATGAW 540
SGLLPTCEAP TESNIEPLVAG LSAAGLSLLT LAPFLWLRL CLRKAKKVPF ASSCQSLESD 600
GSYQKPSYIL

Taxol ProstateA221 DNA SEQUENCE

40
45
Gene name: ESTs; Liprin A2
Unigene number: Hs.306480
Probeset Accession #: N51002
Nucleic Acid Accession #: N51002
Coding sequence: 1-3793 (underlined sequences correspond to start and stop codons)

50
55
60
65
70
75
80
1 11 21 31 41 51
ATGATGTGTG AAGTGATGCC CACGATTAAT GAGGACACCC CAATGAGCCA AAGGGGGTCC 60
CAAGCAGATG GCTCGGACTC AGACTCCCAT TTTGAGCAGC TGATGGTGAA TATGCTAGAT 120
GAAAGGGATC GTCTTCTAGA CACCTCTGGG GAGAGCCAGG AAAGCCTCTC ACTTGCCGAG 180
CAAGACTCTC AGGATGTCTAT CTATGACCGA GACTCACTCC AGAGACAGCT CAATTCAGCC 240
CTGCCACAGG ATATCGAATC OCTAACAGGA GGGCTGGCTG GTTCTAAGGG GGCTGATCCA 300
COGGAATTGG CTGCACTGAC AAAAGAATTA AATGCCCTGCA GGGAGCAACT TCTAGAAAAG 360
GAAGAAGAAA TCTCTGAAC TAAAGCTGAA AGAAACAACA CAAGACTATT ACTGGAGCAT 420
TTGGAGTGCC TTGTGTCAAG ACATGAAGA TCACTAAGAA TGACGGTGGT AAAACGGCAA 480
GCCCGAGTCT CCTCAGGAGT ATCCAGTGAA GTTGAAGTTC TCAAGGCACT GAAATCTTTG 540
TTTGAGCAAC ACAAGGCCCTT GGATGAAAAG GTAAGGGAGC GACTGAGGGT TTCTTTAGAA 600
AGAGTCTCTG CACTGGAAGA AGAACTAGCT GCTGCTAATC AGGAGATGCT TGCTTTGCTT 660
GAACAAAATG TTCTATATCA AAGAAAATG GCATCAAGCG AGGATCCAC AGAGTCAGAA 720
CATCTTGAAG GGATGGAACC TGGACAGAAA GTCCATGAGA AGCGTTTGTG CAATGGTTCT 780
ATAGACTCAA CGATGAALC TAGTCAAATA GTTGAACCTAC AAGAATTGCT TGAAAAGCAA 840
AACTATGAAA TGGCCCGAGT GAAAGAACBT TTAGCAGCCC TTCTTCCCG AGTGGGAGAG 900
GTGGACAGG AAGCAGAGAC AGCAAGAAAG GATCTCATT AAGCAGAGA AATGAACACC 960
AAGTATCAAA GGCACATTAG GGAGGCCATG GCACAAAAGG AAGATATGGA AGAAGAATT 1020
ACRACCCCTG AAAAGCGTTA OCTCAGTGCT CAGAGAGAAT CTACCTCCAT ACATGACATG 1080
AATGATAAAC TAGAAAATGA GTTAGCAAAAT AAGAAGCTA TOCTACGGCA GATGGAAGAG 1140
AAAAACAGAC AGTTACAAGA ACGTCTTGAG CTAGCTGAAC AAAAGTTGCA GCAGACCATG 1200
AGAAAGGCTG AAACCTTGCC TGAAGTAGAG GCTGAACTGG CTCAGAGAA TGCAGCCCTA 1260
ACCAAGGCTG AAGAGAGACA TGGAAATATT GAAGAAGCTA TGAGACATTT AGAGGGTCAA 1320
CTTGAGAGGA AGAATCAAGA ACTTCARAGA GCTAGGCCAA GAGAGAAAT GAATGAGGAG 1380
CATAACAGA GATTATCGGA TACGGTTGAT AGACTTCTGA CTGAATCTCA TGAACGCTA 1440
CAACTACACT TAAAGCAAG AATGGCTGCT CTAGAAGAAA AGAATGTTT AATTCAAGAA 1500
TCAGAAACTT TCAGAAAGAA TCTTGAAGAA TCTTTACATG ATAAGGAAAG ATTGACAGAA 1560
GAAATTGAAA AGCTGAGATC TGAACCTGAC CAATTGAAA TGAGAACTGG CTCCTTAATT 1620
GAACCCACAA TACCAGAAC TCATCTAGAC ACCTCAGCTG AGTTGCGSTA CTCAGTGGGA 1680
TCCCTAGTGG ACRGCCAGTC TGATTACAGA ACACTAAAG TAAATAGAG ACCAAGGAGA 1740
GGCCGCAAG GTGTGCAAG AGATGAGCCA AAGGTGAAT CTCTTGGGGA TCACGAGTGG 1800
AATGAAACTT AACAGATTGG AGTACTAAGC AGCCACCCCT TTGAAGTGA CACTGAAATG 1860
TCTGATATTG TACAGATTGA CAGAGAAACA ATTITTAGCT CAATGGATCT TCTCTCTCA 1920
AGTGGTCATT CGATGCCCA GACGCTAGCC ATGATGCTTC AGGAACAATT GGATGCCATC 1980
AACAAAGAAA TCAGGCTAAT TCAGGAAGAA AAGAATCTA CAGAGTTGCG TGCTGAAGAA 2040
ATTGAAAATA GAGTGGCTAG TGTGAGCTTC GAAGGCCGTA ATTTGGCAAG GGTCCACCCA 2100

5
10
15
20
25
30

```

GGTACCTCCA TTAAGGCTC TGTACAGCT TCATCGCTGG CCAGTTCATC TCCCCCAGT 2160
GGACACTCAA CTCCAAAGCT CACCOCTCGA AGCCTTGCCA GGGAAATGGA TCGGATGGGA 2220
GTATATGACAC TGCCAAAGTGA TCTGAGGAAA CATCGGAGAA AGATTGCACT TGTGGAAGAA 2280
GATGGTCGAG AGGACAAAGC AACAAATTAA TGTGAAACTT CTCTCTCTCC TACCCCTAGA 2340
GCCCTCAGAA TGACTCACAC TCTCCCTTCT TCCTACACAA ATGATGCTCG AAGTAGTTTA 2400
TCTGTCTCTC TTGAGCCAGA AAGCCTCGGG CTGTGCTAGT CCAACAGCAG CCAAGACTCT 2460
CTTCACAAGC CCCCCAAGAA GAAAGGAATC AAGTCTTCAA TAGGACGTTT GTTTGGTAAA 2520
AAAGAAAAGC CTGACTTGG GCAGCTCCGA GGCTTTATGG AGACTGAAGC TGCAGCTCAG 2580
GAGTCCCTGG GCTTAGGCAA ACTCGGAAC CTAGCTGAGA AGGATCGAAG ACTAAAGAAA 2640
AAGCATGAACT TCTCTGAAGA AGCTCGGAGA AAGGGATTAC CTTTGGCCCA GTGGGATGGG 2700
CCAACTGTGG TCGCATGGCT AGAGCTTTGG TTGGGAATGC CTGCGTGGTA CGTGGCAGCC 2760
TGCCGAGCCA ACGTGAAGAG TGCTGCCATC ATGCTCTGCTT TATCTGACAC TGAGATCCAG 2820
AGAGAAATTG GAATCAGCAA TCCACTGCAT CGCTTAAAC TTGATATTAG AATCCAGGAG 2880
ATGGTTTCCC TAACAAGTCC TTCAGCTCCT CCAACATCTC GAACCTCTTC AGGCAACGTT 2940
TGGGTGACTG TGAATTAAGA GGAATAATCT GCAGCTCCAG CAAAAACGAA AGAATCTGAG 3000
GAAGGAAGCT GGGCCCAAGT TCCGCTTTT CTACAGACCC TGGCTTATGG AGATATGAAT 3060
CATGAGTGGG ATGGAATGAA ATGGCTTCCC AGCTTGGGGT TACCTCAGTA CAGAAGTTAC 3120
TTTATGGAAT GCTTGGTAGA TGCAAGAAATG TTAGATCACC TAACAAAAAA AGATCTCCGT 3180
GTCCATTATA AATAGGTGGA TAGTTTCCAT CGAACAAAGT TACAATATGG AATTATGTGC 3240
TTAAAGAGAT TGAATTAAGA CAGAAAGGAA CTAGAAAGAA GACGGGAAGC AAGCCAACAT 3300
GAATATAAAG ACGTGTGTGT GTGGAGCAAT GACCGAATTA TTGCTGGAT ACAAGCAATT 3360
GGACTTGGAG ATGTAACAAA TAATATACCT GAGAGCGGTG TGCATGGCTC ACTTATAGCC 3420
CTGGATGAAA ACTTTGACTA CAGCAGCTTA ACTTATTAT TACAGATTCC AACACAGAAC 3480
ACCCAGGCAA GGCAGATTCT TGAAGAGAAA TACAATAACC TCTTGGCCCT GGGAACTGAA 3540
AGCGGACTGC ATGAAGAAAT TGACAAGAAC TTCAGACGTG GATCAACCTG GAGAAGGCAG 3600
TTTCTCTCTC GTGAAGTACA TGAATCAGC ATGATGCTGG GGTCTCTAGA AACATTACCA 3660
GCTGATTTTA GGTTAACCAAC AACCTCTGG CAATCAAGAA AATGACAAAC AGATGTGCT 3720
TCATCAAGAC TGCAAGAGTT AGACAATCC ACTOTTGCA CATACTCATG TCTCGAGTAA 3780
GCGGCCGCTT TAA

```

A222 Protein sequence:

Gene name: ESTs; Liprin A2
 Unigene number: Hs.306480
 Probeset Accession #: N51002
 Protein Accession #: none found
 Signal sequence: none found
 Transmembrane domains: none found
 AAA domain: 286-539
 SAM domains: 895-964, 1017-1084, 1105-1177
 Cellular Localization: not determined

45
50
55
60
65

```

1 11 21 31 41 51
| | | | |
MMCEVMPTIN EDTPMQRGS QSSGSDSDSH PEQLMVNMLD ERDLRLDTLR ETQESLSLAQ 60
ORLQDVIYDR DSLQRQLNSA LPQDIESLTG GLAGSKGADP PEFAALTKEK NACREQLLEK 120
KEEISELKAE RNNTRLLEH LECIVSRHER SLRMTTVKRO AQSPSGVSGE VEVLLKLSL 180
FERHKAJDEK VRERLRVSLR RVSALEEEEA AANQEIVALR EQNVHTQRKM ASSEGSTESE 240
HLEGMEPQKQ VHEKRLSNGS IDSTDSTSIQ VELQELLEKQ NYEMAQMKER LAALSSRVGE 300
VEQEAETARK DLIKTSEMT KYORDIREAM AQKEDMKERI TTLEKRYLSA QRESTDHDM 360
NDKLELNELAN KEALLQMES KNRQLQERLE LAEQKLQQTM RKAKTLPEVE AELAQRILAL 420
TKASERHONI ESRMRLEGG LEEKNQELQR ARQREKMHES HNKRLSDTV DRLTBSNERL 480
QLHLKERRMAA LSEKNVLQIE SETFRKMLEE SLEHDKERLAE EIEKLSSELD QLMKRTGSLI 540
EPTIPTRLID TSAELRYSVG SLVDSQSDYR TTKVIRPRPR GRMGVRDEP KVKSLGDHEW 600
NETQIGIVLS SHEFFSDTEM SDIDDDIRET LFSSMDLLSP SGHSDAQTLA NMLQEQLDAI 660
NKEIKLIQEE KESTELRAES IENRVASVSL EQLNLARVHP GTSITASVTA SSLASSSPPS 720
GHSTPKLTPR SPAREMDRMG VMILPSDLRK HRRKLAVVEE DGRREDKATIK CSTSPPTPR 780
ALRMTHITPS SYHNDARSSL SVSLFPESLG LGSANSQDS LHKAPKKKGI KSSIGRLFKK 840
KEKARLQQLR GFMETEARAQ ESLGLKLGKT QAEKDRRLKK KHELLEBAR KGLPPAQWDG 900
PTVVAVLELW LGMPAWYVAA CRANVKSGAI MSALSDTEIQ REIGTENELH RLKRLRLATQ 960
MVSITSPSAP PTSRTPSGNV WVTHEEMENL AAPAKTKESE EGSWAQCPVF LQTLAYGDMN 1020
HFWIGNEWLP SLGLPQYRSY FMECLVDARM LDHLTKKDLR VLEKMDSPH RTSLYGIMC 1080
LERLNYDRKE LERRREASQH EIKDVLVSN DRIIRWQAI GLREYANNIL EGGVHGLIA 1140
LDENFDYSSL TLLQLIPTQN TQARQLERE YNNLLALGTE RRLDESDDKH FRGGSTWRRO 1200
FPPREVHGIE MMPGSSSTLP AGFRLTTTSG QSRKMTDVA SSRLQRLDWS TVRTYSCLL

```

A223 DNA SEQUENCE

Gene name: CDA14
 Unigene number: Hs.26813
 Probeset Accession #: N32912
 Nucleic Acid Accession #: NM_016570
 Coding sequence: 1- 1134 (underlined sequences correspond to start and stop codons)

75
80

```

1 11 21 31 41 51
| | | | |
ATGAGGCGAC TGAATCGGAA AAAAACTTTA AGTTTGGTAA AAGAGTTGGA TGCTTTCCG 60
AAGGTTCTCG AGAGCTATGT AGAGACTTCA GCCAGTGGAG GTACAGTTTC TCTAATAGCA 120
TTTACAACTA TGGCTTTATT AACCATATG GAATTTCTCAG TATATCAAGA TACATGGATG 180
AAGTATGAAT ACGAGATAGA CAGGATTTT TCTAGCAAAAT TAAGAAATTA TATAGATATT 240
ACTGTTCGCA TGAAGTGTCA ATATGTTGGA CGCGATGTAT TGGATTTAGC AGAACAATG 300
GTTGCATCGC CAGATGTGTT AGTTTATGAA CCAACAGTAT TTGATCTTTC ACCACAGCAG 360
AAAGAGTGGC AGAGGATGCT GCAGCTGATT CAGAGTAGGC TACAAGAGA GCATTCACCT 420
CAAGATGTGA TATTTAAAG TGCTTTTAAA AGTACATCAA CAGCTCTTCC ACCAAGAGAA 480

```

5
10
GATGATTCAT CACAGTCTCC AAATGCATGC AGAATTCATG GCCATCTATA TGTCAATAAA 540
GTAGCAGGGA ATTTTCACAT AACAGTGGGC AAGGCAATTC CACATCCTCG TGGTCATGCA 600
CATTTGGCAG CACTTGTCAA CCATGAATCT TACAATTTT CTCATAGAAT AGATCATTTG 660
TCITTTGGAG AGCTTGTTC AGCAATTATT AATCCTTTAG ATGGAACCTGA AAAAATTGCT 720
ATAGATCACA ACCAGATGTT CCAATATTTT ATTACAGTTG TGCCCAACAAA ACTACATACA 780
TATAAATAT CAGCAGACAC CCATCAGTTT TCTGTGACAG AAAGGGAACG TATCATTAC 840
CATGCTGCAG GCAGCCATGG AGTCTCTGGG ATATTTATGA AATATGATCT CAGTCTCTTT 900
ATGGTGACAG TTACTGAGGA GCACATGCCA TTCTGGCAGT TTTTGTAAAG ACTCTGTGGT 960
ATTTGTGGAG GAATCTTTTC AACACACAGGC ATGTTACATG GAATTGGAAA ATTTATAGTT 1020
GAAATAATTT GCTGTGCTTT CAGACTTGGG TCCTATAAAC CTGCAATTC TGTTCCTTTT 1080
GAGGATGGCC ACACAGACAA CCACCTACCT CTTTATAGAA ATAATACACA TTGA

15
20
A224 Protein sequence:
Gene name: CDA14
Unigene number: Hs.26813
Probeset Accession #: N32912
Protein Accession #: NP_057654
Signal sequence: none found
Transmembrane domains: none found
Cellular Localization: nuclear

25
30
1 11 21 31 41 51
| | | | |
MRRLNRKKTLLVVKELDAFPVPESEYVETSASGGTVSLIAFTIMALLTIM EFSVYQDTNM 60
KYEYVDKDFSSKLRLNIDI TVAMKQYVGADVLDAETM VASADGLVYEPTVFDLSRQQ 120
KEWQRLQLI QSLRQREHSL QDVIFKSAFK STSTALFPRE DSSSQSPNAC RIHGHLYVNK 180
VAGNFHITVG KALPHFRGEA HLAALVNHEB YNFSHRIDHL SFGEIVPAII NPLDGTAKIA 240
IDHNQMFQYF ITVVPTKLEI YKISADTHQF SVTERERIIN HAAGSRGVSG IFMKYDLSSL 300
MVTVTBEHMP FQGFVRLCG IVGGIFSTIG MLEGIGKFIIV EIICCRFRLG SYKPVNSVVP 360
EDGHTDNHLP LLENWTH

35
40
Uterine
A225 DNA SEQUENCE:
Gene name: ESTs, Weakly similar to JE0350 Anterior gradient-2 [H.sapiens]
Unigene number: Hs.100686
Probeset Accession #: AA487468
Nucleic Acid Accession #: AA487468
Coding sequence: 55-555 (underlined sequences correspond to start and stop codons)

45
50
55
1 11 21 31 41 51
| | | | |
CGGCACCAAG AGCACTGGCC AAGTCAGCTT CTTCTGAGAG AGTCTCTAGA AGACATGATG 60
CTACACTCAG CTTTGGGTCT CAGCCTCTTA CTGCTCACAG TTCTCTCCAA CCTTGCCATT 120
GCAATAAAA AGGAAAGAG AGCTCTCTCAG ACCTCTCAA GAGGATGGGG AGATGACATC 180
ACTTGGGTAC AAACCTATGA AGAAGGTCTC TTTTATGCTC AAAAAGTAA GAAGCCATTA 240
ATGTTTATTC ATCACCCTGA GAATTTCTCA TACTCTCAG CACTAAAGAA AGTATTTGCC 300
CAAAATGAAG AAATACAAGA AATGGCTCAG AATAAGTTCA TCATGCTAAA CCTTATGCAT 360
GAAACCACTG ATAAGAAAT ATCACCCTGAT GGGCAATATG TGCTAGAAAT CATGTTTGT 420
GACCTCTCTT TACAGCTTAG AGCTGACATA GCTGGAAGAT ACTCTAACAG ATTGTACACA 480
TATGAGCCTC GGAATTTACC CCTATTGATA GAAACATGA AGAAGCATT AAGACTTATT 540
CAGTCAGAGC TATTAAGAGAT GATAGAAAA AGCCTTCACT TCAAGAAAGT CAATTTTCAT 600
GAAGAAACC TCTGGCAGAT TGACAAATAC TAAATGTGCA AGTATATAGA TTTTGAATA 660
TTACTATTTA GTTTTTTAA TGTGTTTGA ATAGTCTTAT TAAATAAAT GTTTTTTAA 720
TCTGAAAAA AAAAAAAAAA AAAAAAAAAA

60
65
70
A226 Protein sequence:
Gene name: ESTs, Weakly similar to JE0350 Anterior gradient-2 [H.sapiens]
Unigene number: Hs.100686
Probeset Accession #: AA487468
Protein Accession #: none found
Signal sequence: 1-23
Transmembrane domains: none found
Cellular Localization: secreted

75
1 11 21 31 41 51
| | | | |
MMLHSALGLCLLVTVSSNLATAIKKEKRPQTLSRGWGDITWVQTYEEGLFYAKSEKK 60
PLMVHHLLED CQYSQALKKVFAQNEBIQEM AQNKPIMLNL MHETTDKNLS PDGQYVPRIM 120
FVDPGLTVRA DIAGRYSNRL YTYEPRDLPL LIENMEKALR LIQSEL

80
A227 DNA SEQUENCE
Gene name: G protein-coupled receptor 49 (GPR49) (HG36) (LGR5)
Unigene number: Hs.285529
Probeset Accession #: AA460530
Nucleic Acid Accession #: NM_003667
Coding sequence: 201-2924 (underlined sequences correspond to start and stop codons)

1	11	21	31	41	51	
GTGGCGGCAA	COGGCACTCT	AGTCCCGGCC	GCGCTTCTCC	TCGCGCGCCA	CGCCGTGGGG	60
TCAGGAACCC	GGCGTCTGGC	GCTGCAGAGC	CCCCCTGAGT	TGCAGAGGCC	CACCGAGCGG	120
CGCCCGGGCG	GCCACGGCCC	GTAGCAGTCC	GGTGTCTGCT	TCGCGCGCGG	TCGCGCTCGT	180
GGCCCGCTAC	TTGCGGCACC	ATGGACACCT	CCCGGCTCGG	TGTGCTCTCG	TCCTTGGCTG	240
TGCTGCTGCA	GCTGGCGACC	GGGGGCGACI	CTCCGAGTTC	TGGTGTGTGT	CTGAGGGGCT	300
GCCCCACACA	CTGTCTATGC	GAGCCCGACG	GCAAGATGTT	GCTCAGGGTG	GACTGCTCCG	360
ACCTTGGGGCT	CTCGAGCTGG	CCCTCCACCC	TCAGGCTCTT	CACCTCTTAC	CTAGACCTCA	420
GTATGAACAA	CATCAGTCAG	CTGCTCCCGA	ATCCCTCGCC	CAGTCTCCGC	TTCCCTGGAG	480
AGTTACGCTC	TGCGGGAAAC	GCTCTGACAT	ACATTCOCOA	GGGAGCATTG	ACTGGCCTTT	540
ACAGTCTTAA	AGTCTTTATG	CTGCAGAATA	ATCAGCTAAG	ACACGTACCC	ACAGAGCTTC	600
TGCAGAAATT	GCGAAGCCTT	CAATCCCTGC	GTCTGGATGC	TAACCCATCC	AGCTATGTGC	660
CCCCAAGCTG	TTTCAGTGGC	CTGCATTCCC	TGAGGCACTT	GTGGCTGGAT	GACATGCGGT	720
TAACAGAAAT	CCCCGTCCAG	GCTTTTAGAA	ATTCGAGGCC	ATTGCAAGCC	ATGACCTTGG	780
CCCTGAACAA	AATACACAC	ATACCAAGCT	ATGCTTTTGG	AAACCTCTCC	AGCTTGGTAG	840
TTCTACATCT	CCATAACAAT	AGAATCCACT	CCCTGGGAAA	GAATGCTTTT	GATGGGCTCC	900
ACAGCCTAGA	GACTTTAGAT	TTAAATTACA	ATAACCTTGA	TGAATTCGCC	ACTGCAATTA	960
GGACACTCTC	CAACCTTAAA	GAACCTAGAT	TTCATAGCAA	CAATATCAGG	TCGATACCTG	1020
AGAAAGCATT	TGTAGGCAAC	CCCTCTCTTA	TTACAATACA	TTTCTATGAC	AATCCCATCC	1080
AATTGTGTGG	GAGATCTGCT	TTTCAACATT	TACCTGAAC	AAGAACTCTG	ACTCTGAATG	1140
GTGCTCTTCA	AATAACTGAA	TTTCTGATT	TAACCTGAA	TGCAAACTCT	GAGAGTCTGA	1200
CTTTAACTGG	AGCCACAGATC	TCATCTCTTC	CTCAACCGT	CTGCAATCAG	TTACCTAATC	1260
TCCAAGTCTC	AGATCTGTCT	TACAACTTAT	TAGAAGATT	ACCCAGTTTT	TCAGTCTGCC	1320
AAAGAGCTCA	GAAGATTGAC	CTAAGACATA	ATGAATCTTA	CGAATTAATA	GTGACACTTT	1380
TCAGCAGTTT	GCTTAGCCTC	CGATCGCTGA	ATTTGGCTTG	GAACAAATTT	GCTATATTTC	1440
ACCCCAATGC	ATTTTCCACT	TTGCCATCCC	TAATAAAGCT	GGACCTATCG	TCCAACTCTC	1500
TGTGCTCTTT	TCCTATAACT	GGGTACATG	GTTTAACTCA	CTTAAATTA	ACAGGAAATC	1560
ATGCCCTTCA	GAGCTTGATA	TCATCTGAAA	ACTTTCCAGA	ACTCAAGGTT	ATAGAAATGC	1620
CTTATGCTTA	CCAGTGTCTGT	GCATTTGGAG	TGTGTGAGAA	TGCTATAAG	ATTCTAATC	1680
AATGGAATAA	AGGTGACAA	AGCAGTATGG	ACGACCTTCA	TAAGAAAGAT	GCTGGAATGT	1740
TTACGCTCTA	AGATGAACGT	GACCTTGAAG	ATTTCTGCT	TGACTTTGAG	GAGACCTTGA	1800
AAGCCCTTCA	TTCACTGAGC	TGTTCACTT	CCCGAGGCC	CTTCAACCC	TGTGAACACC	1860
TGCTTGATGG	CTGGCTGATC	AGAAATGGAG	TGTGGACCAT	AGCAGTCTG	GCATCTACTT	1920
GTAATGCTTT	GGTGACTTCA	ACAGTTTTC	GATCCCTCT	GTACATTTCC	CCCATTAAC	1980
TGTTAATTGG	GATCATCGCA	CGATGGAACA	TGCTCAOCCG	AGTCTCCAGT	GCCGTGCTGG	2040
CTGGTGTGGA	TGCTTCACT	TTTGGCAGCT	TTGCAAGACA	TGTTGCCCTG	TGGGAGAATG	2100
GGGTGGTGTG	CCATGTCAAT	GGTTTTTTGT	CCATTTTTGC	TTCAAGATCA	TCTGTTTTCC	2160
TGCTTACTCT	GGCAGCCCTG	GAGCGTGGGT	TCCTGTGAAA	ATATTCTGCA	AAATTTGAAA	2220
CGAAAGCTCC	ATTTTCTAGC	CTGAAAGTAA	TCATTTTGCT	CTGTGCCCTG	CTGGCCTTGA	2280
CCATGGCCGC	AGTTCCTCTG	CTGGGTGGCA	GCAAGTATGG	CGCCTCCCTT	CTCTGCTGCT	2340
CTTTGCTCTT	TGCGGAGCCC	AGCACCATGG	GCTACATGGT	CGCTCTCATC	TTGCTCAATT	2400
CCCTTGTCTT	CCCTATGATG	ACCATTGCC	ACACCAAGCT	CTACTGCAAT	TTGACCAAGG	2460
GAGACCTGGA	GAATATTGG	GACGCTCTTA	TGGTAAACA	CAITGCCCTG	TTGCTCTTCA	2520
CCACCTGATC	CCGTGGCTTT	TCTTGTCTT	CTCCTCTTTA	ATAAAOCTTA	2580	
CATTATATAG	TCCTGAAGTA	ATTAAGTTTA	TCCTTCTGGT	GGTAGTCCCA	CTTCTGCTAT	2640
GTCTCAATAG	CCCTCTCTAC	ATCTGTCTTA	ATCCTCACTT	TAAGGAGGAT	CTGGTGAGCC	2700
TGAGAAAGCA	AACCTACGTC	TGGACAAGAT	CAAAACACCC	AAGCTTGATG	TCAATTAATC	2760
CTGATGATGT	CGAAAAACAG	TCCGTGTGAT	CAACTCAAGC	CTTGGTAACC	TTTACCAGCT	2820
CCAGCATCAC	TTATGACCTG	CCCTCCAGTT	CCGTGCCATC	ACCAGCTTAT	CCAGTGACTG	2880
AGAGCTGCCA	TCCTTCCCTC	GTGGCATTGT	TCCCATGTCT	CTAATTAATA	TGTGAAGGAA	2940
AATGTTTTCA	AGGTGTGAGA	ACCTGAAAAT	GTGAGATTGA	GTATATCAGA	GCAATTAATA	3000
ATAAGAAAG	CTGAGGTGAA	ACTCGGTTTA	AA			

A228 Protein sequence

Gene name: G protein-coupled receptor 49 (GPR49) (HG38) (LGR5)
 Unigene number: Hs.285529
 Protein Accession #: NP_003650.1
 Signal sequence: 1-22
 Transmembrane domains: 557-579, 596-618, 641-663, 687-709, 724-746, 770-792, 803-825
 Cellular Localization: plasma membrane

1	11	21	31	41	51	
MDTSRLGVLL	SLFVLLQLAT	GGSSFRSGVL	LRGCPHCHC	EPDGRMLLRV	DCSDLGLSEL	60
PSNLVFTSY	LDLSMNNISQ	LLPNPLPSLR	FLEELRLAGN	ALTYIPKGAF	TGLYELKVM	120
LQNMQLRHVP	TEALQMLRSL	QSLRLDANHI	SYVFPSCFSG	LESLRLRLD	DNALTEIPVQ	180
APRELALQA	MTLALNKIHH	IPDYAFGNLS	SLVVLHLHNN	RIHSLGKKCF	DGLHSLDTLD	240
LNYNMLDEFP	TAIRTLNLSK	ELGPHSNIR	SIPEKAFVGN	PSLITIRFYD	NPIQFVGRSA	300
FQHLPELRLT	TLNGASQITE	FPDLTGTANL	ESLTLTGAQI	SSLPQTVCNQ	LPNLQVLDLS	360
YNLELDLPSF	SVQCQLQKID	LRHNEIYELK	VDTFQQLSL	RLNLNANNKI	AIIEFNATST	420
LPSLIKLIDL	SNLLSSEFPT	GLHGLTELKL	TGNEALQSLI	SEENFPELKV	IEFPYAYQCC	480
AFGVCEMAYK	ISNQWKNQDN	SSMDLHKDD	AGMFOAQDER	DLDFLIDFE	EDLXALHSVQ	540
CSPSPGFYKP	CEHLLDGLIT	RIGVWTTAVL	ALTCNALVTS	TVFRSPLYIS	PIKLLIGVIA	600
AVNMLTGVS	AVLAGVDAPT	FGSFARHGAW	WENGVGCHVI	GFLSIFASES	SVFLTLAAL	660
ERGFSVKYSA	KFTKAPFSS	LKVIILLCAL	LALTMAAVPL	LGGSKYGASF	LCLPLPFGEF	720
STMGYVVALI	LLNELCFLNM	PLATYTKLYCN	LDRGDLNIN	DCSMVKHIAL	LLFTNCILNC	780
PVALPSFSL	INLTFISPEV	IKFILLVVVP	LPACINPLLY	ILFNPHKED	LVSLRKQTYV	840
WTRSKHPSLM	SINSDVEKQ	SCDSTQALVT	FTSSBITYDL	PPSSVFPSPAY	PVTESCHLSS	900
VAFVPC						

Table 75: See Table 1

Table 76A depicts Seq ID No; UnigeneID; UnigeneTitle; PKey; Predicted Cellular Localization; and Exemplar Accession for each of the sequences in Table 78. The information in Table 76A is linked by SeqID No to Table 78.

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

Seq ID No:	Pkey:	ExAccn:	UnigeneID:	Unigene Title:	Pred Subcell Loc:
Seq ID No:	Sequence ID No for sequences in table				
Pkey:	Unique Eos probeset identifier number				
ExAccn:	Exemplar Accession number, Genbank accession number				
UnigeneID:	Unigene number				
Unigene Title:	Unigene gene title				
Pred Subcell Loc:	Predicted sub-cellular localization				
Seq ID No	Pkey	ExAccn	UnigeneID	Unigene Title	Pred Subcell Loc
Seq ID 1 & 2	425023	AW956889	Hs.154210	endothelial differentiation, sphingolipid	plasma membrane
Seq ID 3 & 4	424503	NM_002205	Hs.149609	Integrin, alpha 5 (fibronectin receptor,	plasma membrane
Seq ID 5 & 6	429423	A016712	Hs.287797	Integrin, beta 1 (fibronectin receptor,	plasma membrane
Seq ID 7 & 8	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin	secreted
Seq ID 9 & 10	419172	AW338625	Hs.22120	ESTs; similar to TRANSMEMBRANE 4 SUPERF	plasma membrane
Seq ID 11 & 12	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	secreted
Seq ID 13 & 14	418007	M13509	Hs.83169	matrix metalloproteinase 1 (interstitial	secreted
Seq ID 15 & 16	407836	T79340	Hs.200272	B-cell CLL/lymphoma 6, member B, zinc fi	intracell
Seq ID 17 & 18	414577	A056548	Hs.72116	hypothetical protein FLJ20992 similar to	secreted
Seq ID 19 & 20	418738	AW388633	Hs.6682	solute carrier family 7, (cationic amino	plasma membrane
Seq ID 21 & 22	428368	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin	secreted
Seq ID 23 & 24	415138	C18356	Hs.295944	tissue factor pathway inhibitor 2	secreted
Seq ID 25 & 26	429276	AF056085	Hs.198612	G protein-coupled receptor 51	plasma membrane
Seq ID 27 & 28	418994	AA296520	Hs.89546	selectin E (endothelial adhesion molecu	plasma membrane
Seq ID 29 & 30	407975	X89426	Hs.41716	endothelial cell-specific molecule 1	secreted
Seq ID 31 & 32	429113	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p	intracell
Seq ID 33 & 34	418606	AA084248	Hs.85339	G protein-coupled receptor 39	plasma membrane
Seq ID 35 & 36	429961	D13666	Hs.136348	perforin (OSF-2os)	secreted
Seq ID 37 & 38	414812	X72755	Hs.77367	monokine induced by gamma interferon	secreted
Seq ID 39 & 40	417433	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein	plasma membrane
Seq ID 41 & 42	417433	BE270266	Hs.82128	ST4 oncofetal trophoblast glycoprotein	plasma membrane
Seq ID 43 & 44	424399	A1906887		A1906887/IL-BT095-190199-019 BT095 Homo	secreted
Seq ID 45 & 46	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (po	secreted
Seq ID 47 & 48	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	secreted
Seq ID 49 & 50	444381	BE387335	Hs.283713	ESTs, Weakly similar to S64054 hypotheti	secreted
Seq ID 51 & 52	439589	AW602186	Hs.222399	CEGP1 protein	secreted
Seq ID 53 & 54	411558	AA102570	Hs.70725	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
Seq ID 55 & 56	400303	AA242758	Hs.78138	LIV-1 protein, estrogen regulated	plasma membrane
Seq ID 57 & 58	411789	AF245505	Hs.72157	Adiccan	secreted
Seq ID 59 & 60	428698	AA852773	Hs.334838	KIAA1868 protein	plasma membrane
Seq ID 61 & 62	450098	W27249	Hs.8109	hypothetical protein FLJ21080	intracell
Seq ID 63 & 64	421562	AF026692	Hs.105700	secreted frizzled-related protein 4	secreted
Seq ID 65 & 66	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR	plasma membrane
Seq ID 67 & 68	415539	A1733881	Hs.72472	BMP-R1B	plasma membrane
Seq ID 69 & 70	418636	N32536	Hs.42645	solute carrier family 18 (monocarboxylic	plasma membrane
Seq ID 71 & 72	418636	N32536	Hs.42645	solute carrier family 16 (monocarboxylic	plasma membrane
Seq ID 73 & 74	409079	W87707	Hs.82065	Interleukin 6 signal transducer (gp130,	plasma membrane
Seq ID 75 & 76	442082	R41823	Hs.7413	ESTs	plasma membrane
Seq ID 77 & 78	400297	A1127076	Hs.306201	hypothetical protein DKFZp564O1278	plasma membrane
Seq ID 79 & 80	451398	A1793124	Hs.144479	ESTs	intracell
Seq ID 81 & 82	428220	AW207206		ESTs	plasma membrane
Seq ID 83 & 84	421524	AA312082	Hs.105445	GDNF family receptor alpha 1	plasma membrane
Seq ID 85 & 86	423242	AL039402	Hs.125783	DEME-6 protein	plasma membrane
Seq ID 87 & 88	423242	AL039402	Hs.125783	DEME-6 protein	intracell
Seq ID 89 & 90	423242	AL039402	Hs.125783	DEME-6 protein	intracell
Seq ID 91 & 92	452190	H26735	Hs.91668	Homo sapiens clone PP149B unknown mRNA	intracell
Seq ID 93 & 94	462190	H26735	Hs.91668	Homo sapiens clone PP149B unknown mRNA	intracell
Seq ID 95 & 96	325372			Phase 2 & 3 Exons	cytoplasmic
Seq ID 97 & 98	450375	AA008847		alpha disintegrin and metalloproteinase doma	plasma membrane
Seq ID 99 & 100	428245	AW963419	Hs.155223	stanniocalcin 2	secreted
Seq ID 101 & 102	425247	NM_005940	Hs.155324	matrix metalloproteinase 11 (stromelysin	secreted
Seq ID 103 & 104	429353	AL117406	Hs.335891	ATP-binding cassette transporter MRP8	plasma membrane
Seq ID 105 & 106	429353	AL117406	Hs.335891	ATP-binding cassette transporter MRP8	plasma membrane
Seq ID 107 & 108	432201	A1538613	Hs.288241	Transmembrane protease, serine 3	plasma membrane
Seq ID 109 & 110	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	secreted
Seq ID 111 & 112	446163	AA026880	Hs.25252	prolactin receptor	plasma membrane
Seq ID 113 & 114	442117	AW664954	Hs.128899	ESTs; hypothetical protein for IMAGE:447	plasma membrane
Seq ID 115 & 116	428179	A1127772	Hs.273696	serum/glucocorticoid regulated kinase-II	intracell
Seq ID 117 & 118	431211	M86849	Hs.323733	gap junction protein, beta 2, 26kD (conn	plasma membrane
Seq ID 119 & 120	447033	A1357412	Hs.157601	ESTs	secreted
Seq ID 121 & 122	447033	A1357412	Hs.157601	ESTs	secreted
Seq ID 123 & 124	447033	A1357412	Hs.157601	ESTs	secreted
Seq ID 125 & 126	115522	BE614387	Hs.333893	c-Myc target JPO1	intracell
Seq ID 127 & 128	452579	Z42387	Hs.83983	transmembrane, prostate androgen induced	plasma membrane
Seq ID 129 & 130	446051	BE048061	Hs.37054	ephrin-A3	plasma membrane
Seq ID 131 & 132	422048	NM_012445	Hs.288126	spondin 2, extracellular matrix protein	secreted
Seq ID 133 & 134	410418	D31382	Hs.63325	transmembrane protease, serine 4	plasma membrane
Seq ID 135 & 136	446342	BE298665	Hs.14846	Homo sapiens mRNA; cDNA DKFZp564D016 (fr	plasma membrane
Seq ID 137 & 138	422260	AA315993	Hs.105484	regenerating gene type IV	secreted

	Seq ID 139 & 140	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	secreted
	Seq ID 141 & 142	409041	AB033025	Hs.50081	Hypothetical protein, XP_051860 (KIAA119	secreted
	Seq ID 143 & 144	111929	AF027208	Hs.112360	prominin (mouse)-like 1	plasma membrane
5	Seq ID 145 & 146	111929	AF027208	Hs.112360	prominin (mouse)-like 1	plasma membrane
	Seq ID 147 & 148	111929	AF027208	Hs.112360	prominin (mouse)-like 1	plasma membrane
	Seq ID 149 & 150	104888	AW939591	Hs.5940	murin 13, epithelial transmembrane	plasma membrane
	Seq ID 151 & 152	420159	AJ572490	Hs.99786	Homo sapiens cDNA: FLJ21245 fis, clone C	plasma membrane
	Seq ID 153 & 154	422330	D30783	Hs.115263	epiregulin	plasma membrane
10	Seq ID 155 & 156	452461	N78223	Hs.108106	transcription factor	intracell
	Seq ID 157 & 158	413324	V00571	Hs.75294	corticotropin releasing hormone	secreted
	Seq ID 159 & 160	412420	AL035668	Hs.73853	bone morphogenetic protein 2	secreted
	Seq ID 161 & 162	416658	U03272	Hs.79432	fibillin 2 (congenital contractural ara	secreted
	Seq ID 163 & 164	407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antag	secreted
15	Seq ID 165	402230			Fgenesh predicted: CYTOCHROME P450 4F5 (intracell
	Seq ID 166 & 167	402230			Fgenesh predicted: CYTOCHROME P450 4F5 (intracell
	Seq ID 168 & 169	432829	W60377	Hs.57772	ESTs	intracell
	Seq ID 170 & 171	425721	AC002115	Hs.159309	uropod 1A	plasma membrane
	Seq ID 172 & 173	420370	Y13645	Hs.97234	uropod 2	plasma membrane
20	Seq ID 174 & 175	437852	BE001836	Hs.256897	ESTs, Weakly similar to DJ365012.1 [H.s	plasma membrane
	Seq ID 176 & 177	402075			ENSP00000251058; Plasma membrane calcium	secreted
	Seq ID 178 & 179	421110	AJ250717	Hs.1355	cathepsin E	secreted
	Seq ID 180 & 181	451668	Z43948	Hs.326444	cartilage acidic protein 1	plasma membrane
	Seq ID 182 & 183	451668	Z43948	Hs.326444	cartilage acidic protein 1	secreted
25	Seq ID 184 & 185	451668	Z43948	Hs.326444	cartilage acidic protein 1	intracell
	Seq ID 186 & 187	408243	Y00787	Hs.624	Interleukin B	secreted
	Seq ID 188 & 189	422282	AF019225	Hs.114309	apolipoprotein L	secreted
	Seq ID 190 & 191	425852	AK001504	Hs.159651	death receptor 6, TNF superfamily member	plasma membrane
	Seq ID 192 & 193	439738	BE246502	Hs.9598	sema domain, immunoglobulin domain (Ig),	plasma membrane
30	Seq ID 194 & 195	404875			NM_022819; Homo sapiens phospholipase A2	intracell
	Seq ID 196 & 197	425883	AL137708	Hs.161031	Homo sapiens mRNA; cDNA DKFZp434K0322 (I	plasma membrane
	Seq ID 198 & 199	404977			Insulin-like growth factor 2 (somatomedi	secreted
	Seq ID 200 & 201	420876	AA918425	Hs.177744	ESTs	plasma membrane
	Seq ID 202 & 203	427747	AW411425	Hs.180655	serine/threonine kinase 12	intracell
35	Seq ID 204 & 205	420281	AJ623693	Hs.323494	Predicted cation efflux pump	plasma membrane
	Seq ID 206 & 207	446673	NM_018361	Hs.15871	LPAP for lysophosphatidic acid phosphata	intracell
	Seq ID 208	437553	AB209935	Hs.130497	ESTs, Weakly similar to MATB_HUMAN CHLOR	plasma membrane
	Seq ID 209 & 210	437553	AB209935	Hs.130497	ESTs, Weakly similar to MATB_HUMAN CHLOR	plasma membrane
	Seq ID 211 & 212	437553	AB209935	Hs.130497	ESTs, Weakly similar to MATB_HUMAN CHLOR	plasma membrane
40	Seq ID 213 & 214	426900	AW163564	Hs.142375	ESTs	plasma membrane
	Seq ID 215 & 216	426900	AW163564	Hs.142375	ESTs	plasma membrane
	Seq ID 217 & 218	426900	AW163564	Hs.142375	ESTs	plasma membrane
	Seq ID 219 & 220	405932			C15000305.gi33806122.gbjAAC69198.1 (AF0	intracell
	Seq ID 221 & 222	405932			C15000305.gi33806122.gbjAAC69198.1 (AF0	intracell
45	Seq ID 223 & 224	424008	R02740	Hs.137555	putative chemokine receptor; GTP-binding	plasma membrane
	Seq ID 225 & 226	424008	R02740	Hs.137555	putative chemokine receptor; GTP-binding	plasma membrane
	Seq ID 227 & 228	424008	R02740	Hs.137555	putative chemokine receptor; GTP-binding	plasma membrane
	Seq ID 229 & 230	444342	NM_014398	Hs.10887	similar to lysosome-associated membrane	secreted
	Seq ID 231 & 232	412179	Y15221	Hs.103982	small inducible cytokine subfamily 6 (Cy	secreted
50	Seq ID 233 & 234	417079	U85690	Hs.81134	Interleukin 1 receptor antagonist	secreted
	Seq ID 235 & 236	412986	X81120	Hs.75110	cannabinoid receptor 1 (brain)	plasma membrane
	Seq ID 237 & 238	412986	X81120	Hs.75110	cannabinoid receptor 1 (brain)	plasma membrane
	Seq ID 239 & 240	412986	X81120	Hs.75110	cannabinoid receptor 1 (brain)	plasma membrane
	Seq ID 241 & 242	447072	D61594	Hs.17279	tyrosylprotein sulfotransferase 1	intracell
55	Seq ID 243 & 244	419723	AL120193	Hs.339810	longevity assurance (LAG1, S. cerevisiae	secreted
	Seq ID 245 & 246	419723	AL120193	Hs.339810	longevity assurance (LAG1, S. cerevisiae	plasma membrane
	Seq ID 247 & 248	430880	X54232	Hs.2699	glypican 1	plasma membrane
	Seq ID 249 & 250	456759	BE269150	Hs.127792	delta (Drosophila)-like 3	plasma membrane
	Seq ID 251	429466	M85835	Hs.12827	ESTs	
60	Seq ID 252	429466	M85835	Hs.12827	ESTs	
	Seq ID 253 & 254	419721	NM_001650	Hs.288850	aquaporin 4	plasma membrane
	Seq ID 255 & 256	407034	U84540		gb:Human dystrobrevin isoform DTN-3 (DTN	secreted
	Seq ID 257 & 258	413472	BE242870	Hs.75379	solute carrier family 1 (glial high affi	plasma membrane
	Seq ID 259 & 260	438380	T06430	Hs.6194	chondroitin sulfate proteoglycan BEHAB/b	secreted & plasma membrane
65	Seq ID 261 & 262	428271	AF028547	Hs.168047	chondroitin sulfate proteoglycan 3 (neur	secreted
	Seq ID 263 & 264	419704	AA428104	Hs.45057	ESTs	intracell
	Seq ID 265 & 266	444471	AB020684	Hs.11217	KIAA0877 protein	plasma membrane
	Seq ID 267 & 268	409395	U46745	Hs.336678	dystrobrevin, alpha	secreted
	Seq ID 269 & 270	413063	AL035737	Hs.76184	chitinase 3-like 1 (cartilage glycoprote	secreted
70	Seq ID 271 & 272	433000	AI034361	Hs.135160	lung type-I cell membrane-associated gly	plasma membrane
	Seq ID 273	458435	AI418718	Hs.144121	ESTs, Weakly similar to T46916 hypotheti	
	Seq ID 274	458435	AI418718	Hs.144121	ESTs, Weakly similar to T46916 hypotheti	
	Seq ID 275 & 276	424343	AW956360	Hs.4748	adenylate cyclase activating polypeptide	plasma membrane
	Seq ID 277 & 278	424898	U58515	Hs.154138	chitinase 3-like 2	secreted
75	Seq ID 279 & 280	412709	AL022327	Hs.74518	KIAA0027 protein	plasma membrane
	Seq ID 281 & 282	435615	Y15065	Hs.4975	potassium voltage-gated channel, KCQT-IIk	plasma membrane
	Seq ID 283 & 284	404049			NM_018937; Homo sapiens protocadherin bc	plasma membrane
	Seq ID 285 & 286	418932	L34059	Hs.89484	cadherin 4, type 1, R-cadherin (retinal)	plasma membrane
	Seq ID 287 & 288	404029			NM_018935; Homo sapiens protocadherin bc	plasma membrane
	Seq ID 289 & 290	436480	AJ271843	Hs.87469	putative acid-sensing ion channel	intracell
80	Seq ID 291 & 292	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	secreted
	Seq ID 293 & 294	452401	NM_007115	Hs.29352	tumor necrosis factor, alpha-induced pro	secreted
	Seq ID 295 & 296	436895	AF037335	Hs.5338	carbonic anhydrase XII	plasma membrane
	Seq ID 297 & 298	421471	U90545	Hs.327179	solute carrier family 17 (sodium phospho	plasma membrane

5	Seq ID 299 & 300	428296	NM_003058	Hs.183672	solute carrier family 22 (organic cation	plasma membrane
	Seq ID 301 & 302	423508	AW604297	Hs.129711	hepatitis A virus cellular receptor 1	plasma membrane
	Seq ID 303 & 304	450001	NM_001044	Hs.406	solute carrier family 6 (neurotransmitter)	plasma membrane
	Seq ID 305 & 306	410407	X66839	Hs.63287	carbonic anhydrase IX	plasma membrane
	Seq ID 307 & 308	453496	AA442103	Hs.33084	solute carrier family 2 (facilitated glu	plasma membrane
10	Seq ID 309 & 310	420737	L08096	Hs.99899	CD70 ; tumor necrosis factor (ligand) s	plasma membrane
	Seq ID 311 & 312	309931	AW341683		gbc:hd13d01.x1 Soares_NFL_T_GBC_S1 Homo s	plasma membrane
	Seq ID 313 & 314	412719	AW016610	Hs.816	ESTs	intracell
	Seq ID 315 & 316	417034	NM_006183	Hs.80962	neurotensin	secreted
	Seq ID 317 & 318	430486	BE062109	Hs.241551	chloride channel, calcium activated, fam	plasma membrane
15	Seq ID 319 & 320	413753	U17760	Hs.75517	laminin, beta 3 (nicotin (125kD), kallini	secreted
	Seq ID 321 & 322	425850	NM_001944	Hs.1925	desmoglein 3 (pemphigus vulgaris antigen	plasma membrane
	Seq ID 323 & 324	423673	BE003054	Hs.1695	matrix metalloproteinase 12 (macrophage	secreted
	Seq ID 325 & 326	418663	AK001100	Hs.41680	desmocollin 3	plasma membrane
	Seq ID 327 & 328	418663	AK001100	Hs.41690	desmocollin 3	plasma membrane
20	Seq ID 329 & 330	429610	AB024937	Hs.211092	LUNX protein; PLUNC (palate lung and nas	secreted
	Seq ID 331 & 332	406590	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	plasma membrane
	Seq ID 333 & 334	431846	BE019924	Hs.271580	uroplakin 1B	plasma membrane
	Seq ID 335 & 336	422158	L10343	Hs.112341	protease inhibitor 3, skin-derived (SKAL	secreted
	Seq ID 337 & 338	431958	X83629	Hs.2877	cadherin 3, type 1, P-cadherin (placenta	plasma membrane
25	Seq ID 339 & 340	437044	AL036864	Hs.69517	differentially expressed in Fanconi's en	plasma membrane
	Seq ID 341 & 342	428484	AF104032	Hs.184801	solute carrier family 7 (cationic amino	plasma membrane
	Seq ID 343 & 344	429211	AF052693	Hs.198249	gap junction protein, beta 5 (connexin 3	plasma membrane
	Seq ID 345 & 346	417389	BE260964	Hs.82945	midline (neurite growth-promoting factor	secreted
	Seq ID 347 & 348	431009	BE149762	Hs.48956	gap junction protein, beta 6 (connexin 3	plasma membrane
30	Seq ID 349 & 350	417542	J04129	Hs.82269	prostaglandin-associated endometrial prote	secreted
	Seq ID 351 & 352	449230	BE613348	Hs.211579	melanoma cell adhesion molecule	plasma membrane
	Seq ID 353 & 354	410555	U92849	Hs.64311	a disintegrin and metalloproteinase doma	plasma membrane
	Seq ID 355 & 356	410555	U92849	Hs.64311	a disintegrin and metalloproteinase doma	plasma membrane
	Seq ID 357 & 358	424587	J05070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	secreted
35	Seq ID 359 & 360	418482	BE001598	Hs.85262	Integrin, beta 4	plasma membrane
	Seq ID 361 & 362	410274	AA381807	Hs.61762	hypoxia-inducible protein 2	secreted
	Seq ID 363 & 364	439606	W79123	Hs.58561	G protein-coupled receptor 87	plasma membrane
	Seq ID 365 & 366	404877			NM_005365:Homo sapiens melanoma antigen,	intracell
	Seq ID 367 & 368	444781	NM_014400	Hs.11950	GPI-anchored metastasis-associated prote	plasma membrane
40	Seq ID 369 & 370	109424	NM_005329	Hs.85952	hyaluronan synthase 3	plasma membrane
	Seq ID 371 & 372	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
	Seq ID 373 & 374	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
	Seq ID 375 & 376	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
	Seq ID 377 & 378	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
45	Seq ID 379 & 380	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
	Seq ID 381 & 382	415817	U88967	Hs.78867	protein tyrosine phosphatase, receptor-t	plasma membrane
	Seq ID 383 & 384	421817	AF146074	Hs.108560	ATP-binding cassette, sub-family C (CFTR	plasma membrane
	Seq ID 385 & 386	418678	NM_001327	Hs.87225	cancer/testis antigen (NY-ESO-1)	intracell
	Seq ID 387 & 388	418678	NM_001327	Hs.87225	cancer/testis antigen (NY-ESO-1)	intracell
50	Seq ID 389 & 390	409420	Z15008	Hs.54451	laminin, gamma 2 (nicotin (100kD), kalini	secreted
	Seq ID 391 & 392	332180	AF134160	Hs.7327	claudin 1	plasma membrane
	Seq ID 393 & 394	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	plasma membrane
	Seq ID 395 & 396	408790	AW580227	Hs.47860	neurotrophic tyrosine kinase, receptor,	plasma membrane
	Seq ID 397 & 398	439223	AW236299	Hs.250518	UL16 binding protein 2	plasma membrane
55	Seq ID 399 & 400	409757	NM_001898	Hs.123114	cystatin SN	secreted
	Seq ID 401 & 402	428969	AF120274	Hs.194689	artemin	secreted
	Seq ID 403 & 404	428969	AF120274	Hs.194689	artemin	secreted
	Seq ID 405 & 406	428969	AF120274	Hs.194689	artemin	secreted
	Seq ID 407 & 408	428969	AF120274	Hs.194689	artemin	secreted
60	Seq ID 409 & 410	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	secreted
	Seq ID 411 & 412	450701	H39960	Hs.288467	hypothetical protein XP_098151 (leucine-	secreted
	Seq ID 413 & 414	414774	X02419	Hs.77274	plasminogen activator, urokinase	secreted
	Seq ID 415 & 416	407944	R34008	Hs.239727	desmocollin 2	plasma membrane
	Seq ID 417 & 418	407944	R34008	Hs.239727	desmocollin 2	plasma membrane
65	Seq ID 419 & 420	428486	AW583497	Hs.184604	pancreatic polypeptide	secreted
	Seq ID 421 & 422	457489	AI693815	Hs.127179	cryptic gene	secreted
	Seq ID 423 & 424	432874	W94322	Hs.279651	melanoma inhibitory activity	secreted
	Seq ID 425 & 426	445891	AW391342	Hs.199460	DPCR1 protein	plasma membrane
	Seq ID 427 & 428	445891	AW391342	Hs.199460	DPCR1 protein	plasma membrane
70	Seq ID 429 & 430	404682			C9001188:g[12738842(ref)]NP_073725.1] p	secreted
	Seq ID 431 & 432	429547	AW009186	Hs.99376	ESTs	secreted
	Seq ID 433 & 434	429521	NM_007231	Hs.162211	solute carrier family 6 (neurotransmitter	plasma membrane
	Seq ID 435 & 436	407242	M18728		gbc:Human nonspecific crossreacting antig	plasma membrane
	Seq ID 437 & 438	407242	M18728		gbc:Human nonspecific crossreacting antig	plasma membrane
75	Seq ID 439 & 440	407242	M18728		gbc:Human nonspecific crossreacting antig	plasma membrane
	Seq ID 441 & 442	432598	AJ224741	Hs.278461	matrilin 3	secreted
	Seq ID 443 & 444	444006	BE395085	Hs.10086	type 1 transmembrane protein Fn14	plasma membrane
	Seq ID 445 & 446	423685	BE350494	Hs.49753	uveal autoantigen with coiled coil domai	intracell
	Seq ID 447 & 448	423692	H10233	Hs.2265	secretory granule, neuroendocrine protei	secreted
80	Seq ID 449 & 450	423697	NM_003816	Hs.2442	a disintegrin and metalloproteinase doma	plasma membrane
	Seq ID 451 & 452	448030	N30714	Hs.325960	membrane-spanning 4-domains, subfamily A	plasma membrane
	Seq ID 453 & 454	422109	S73265	Hs.1473	gastrin-releasing peptide	secreted
	Seq ID 455 & 456	419236	AW470411	Hs.288433	neurotrophin	plasma membrane
	Seq ID 457 & 458	448048	Z45051	Hs.22920	similar to S68401 (cattle) glucose induc	plasma membrane
	Seq ID 459 & 460	427333	AF087797	Hs.176658	aquaporin 8	secreted
	Seq ID 461 & 462	417931	W95642	Hs.82961	trefoil factor 3 (intestinal)	secreted
	Seq ID 463 & 464	419216	AU076718	Hs.164021	small inducible cytokine subfamily B (Cy	secreted

5	Seq ID 465 & 466	431629	AU077025	Hs.265827	Interferon, alpha-inducible protein (c)	secreted
	Seq ID 467 & 468	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	secreted
	Seq ID 469 & 470	452194	AI694413	Hs.332649	Ubiquitin-like protein FAT10777	plasma membrane
	Seq ID 471 & 472	452194	AI694413	Hs.332649	Ubiquitin-like protein FAT10777	plasma membrane
	Seq ID 473 & 474	426322	J05068	Hs.2012	transcobalamin 1 (vitamin B12 binding pr	secreted
10	Seq ID 475 & 476	429010	Y18198	Hs.194725	one cut domain, family member 2	intracell
	Seq ID 477 & 478	431462	AW583672	Hs.256311	granin-like neuroendocrine peptide precu	secreted
	Seq ID 479 & 480	448243	AW369771	Hs.52620	Integrin, beta 8	plasma membrane
	Seq ID 481 & 482	426427	M86699	Hs.169840	TTK protein kinase	Intracell
	Seq ID 483 & 484	428187	AI687303	Hs.285529	G protein-coupled receptor 49	plasma membrane
15	Seq ID 485 & 486	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
	Seq ID 487 & 488	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
	Seq ID 489 & 490	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
	Seq ID 491 & 492	428579	NM_005756	Hs.184942	G protein-coupled receptor 64	plasma membrane
	Seq ID 493 & 494	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6	secreted
20	Seq ID 495 & 496	422278	AF072873	Hs.114218	frizzled (Drosophila) homolog 6	plasma membrane
	Seq ID 497 & 498	424620	AA101043	Hs.151264	kallikrein 7 (chymotryptic, stratum com	secreted
	Seq ID 499 & 500	408400			kallikrein 8 (neurosin/kovasin) (KLK8)	secreted
	Seq ID 501 & 502	431130	NM_006103	Hs.2719	HE4; epididymis-specific, whey-acidic pr	secreted
	Seq ID 503 & 504	420440	NM_002407	Hs.97644	mammaglobin 2	secreted
25	Seq ID 505 & 506	428450	NM_014791	Hs.184339	KIAA0175 gene product	intracell
	Seq ID 507 & 508	446619	AU076643	Hs.313	secreted phosphoprotein 1 (osteopontin,	secreted
	Seq ID 509 & 510	436982	AB018305	Hs.5378	spondin 1, (f-spondin) extracellular mat	secreted
	Seq ID 511 & 512	453392	U23752	Hs.32964	SRV (sex determining region Y)-box 11	secreted
	Seq ID 513	431989	AW972870	Hs.291069	ESTs	Intracell
30	Seq ID 514	439820	AL360204	Hs.283853	Homo sapiens mRNA full length insert cDN	
	Seq ID 515 & 516	409178	BE383948	Hs.50915	kallikrein 5	secreted
	Seq ID 517 & 518	428514	BE618633	Hs.170195	bone morphogenetic protein 7 (osteogenic	secreted
	Seq ID 519 & 520	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	plasma membrane
	Seq ID 521 & 522	421478	AI683243	Hs.97258	ESTs, Moderately similar to S29539 ribos	plasma membrane
35	Seq ID 523 & 524	425776	U25128	Hs.159499	parathyroid hormone receptor 2	plasma membrane
	Seq ID 525 & 526	425776	U25128	Hs.159499	parathyroid hormone receptor 2	plasma membrane
	Seq ID 527 & 528	452097	AB002364	Hs.27916	a disintegrin-like and metalloprotease (secreted
	Seq ID 529 & 530	416530	U62801	Hs.79361	kallikrein 6 (neurosin, zyme)	secreted
	Seq ID 531 & 532	431515	NM_012162	Hs.258583	endothelial differentiation, lysophospha	plasma membrane
40	Seq ID 533 & 534	419462	U33635	Hs.90572	PTK7 protein tyrosine kinase 7	plasma membrane
	Seq ID 535 & 536	412078	X69699	Hs.73149	paired box gene 8	intracell
	Seq ID 537 & 538	412078	X69699	Hs.73149	paired box gene 8	intracell
	Seq ID 539 & 540	411773	NM_006799	Hs.72026	protease, serine, 21 (testisin)	secreted
	Seq ID 541 & 542	407792	AI077715	Hs.39384	putative secreted ligand homologous to f	secreted
45	Seq ID 543 & 544	431616	AA508652	Hs.195839	ESTs, Weakly similar to I38022 hypotheti	plasma membrane
	Seq ID 545 & 546	452792	AB037765	Hs.30852	KIAA1344 protein	plasma membrane
	Seq ID 547 & 548	400294	N95786	Hs.278696	Homo sapiens protein mRNA, complete cds	plasma membrane
	Seq ID 549 & 550	432653	N62086	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
	Seq ID 551 & 552	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
50	Seq ID 553 & 554	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
	Seq ID 555 & 556	432653	N62096	Hs.293185	ESTs, Weakly similar to JC7328 amino aci	plasma membrane
	Seq ID 557 & 558	400290	H18836	Hs.31608	hypothetical protein FLJ20041	plasma membrane
	Seq ID 559 & 560	410001	AB041036	Hs.57771	kallikrein 11	secreted
	Seq ID 561 & 562	418396	A765805	Hs.26691	ESTs	plasma membrane
55	Seq ID 563 & 564	451027	AW519204	Hs.40808	ESTs	plasma membrane
	Seq ID 565 & 566	446057	AM20227	Hs.149358	ESTs, Weakly similar to A46010 X-linked	plasma membrane
	Seq ID 567 & 568	433466	AA508353	Hs.105314	relaxin 1 (H1)	secreted
	Seq ID 569 & 570	453370	AI470523	Hs.139335	ATP-binding cassette, sub-family C (CFTR	plasma membrane
	Seq ID 571 & 572	453370	AI470523	Hs.139335	ATP-binding cassette, sub-family C (CFTR	plasma membrane
60	Seq ID 573 & 574	414569	AF109298	Hs.118258	prostate cancer associated protein 1	plasma membrane
	Seq ID 575 & 576	413435	X51405	Hs.76360	carboxypeptidase E	secreted
	Seq ID 577 & 578	426501	AW043782	Hs.293816	ESTs	plasma membrane
	Seq ID 579 & 580	448099	AF178274	Hs.22791	transmembrane protein with EGF-like and	plasma membrane
	Seq ID 581 & 582	408369	R38438	Hs.182575	soluble carrier family 15 (H777 transport	plasma membrane
65	Seq ID 583 & 584	412628	AI972402	Hs.308051	hypothetical protein MGC2648	secreted
	Seq ID 585 & 586	403047			NM_005656: Homo sapiens transmembrane pr	plasma membrane
	Seq ID 587 & 588	403047			NM_005656: Homo sapiens transmembrane pr	plasma membrane
	Seq ID 589 & 590	408430	S79876	Hs.44828	diacylglycerol lipase IV (CD28, adenosine	plasma membrane
	Seq ID 591 & 592	445413	AA151342	Hs.12677	CGI-147 protein	secreted
70	Seq ID 593 & 594	451982	F13036	Hs.27373	Homo sapiens mRNA; cDNA DKFZp564Q1763 (f	intracell
	Seq ID 595 & 596	427958	AA418000	Hs.98280	potassium intermediate/small conductance	plasma membrane
	Seq ID 597 & 598	421887	AW181450	Hs.109201	CGI-88 protein	plasma membrane
	Seq ID 599 & 600	425071	NM_013889	Hs.154424	deiodinase, iodothyronine, type II	secreted
	Seq ID 601 & 602	432101	AI918950	Hs.123642	EphA3	plasma membrane
75	Seq ID 603 & 604	407786	AA687538	Hs.38972	tetraspan 1	plasma membrane
	Seq ID 605 & 606	416836	D54745	Hs.80247	cholecystokinin	secreted
	Seq ID 607 & 608	416539	Y07909	Hs.79368	epithelial membrane protein 1	plasma membrane
	Seq ID 609 & 610	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
	Seq ID 611 & 612	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
80	Seq ID 613 & 614	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
	Seq ID 615 & 616	131083	Y09763	Hs.22785	gamma-aminobutyric acid (GABA) A recepto	plasma membrane
	Seq ID 617 & 618	422424	AI86431	Hs.298638	prostate differentiation factor	secreted
	Seq ID 619 & 620	428970	BE276891	Hs.194691	retinoic acid induced 3	plasma membrane
	Seq ID 621 & 622	428330	L22524	Hs.2256	matrix metalloproteinase 7 (matrilysin,	secreted
	Seq ID 623 & 624	430018	AW300887	Hs.26638	membrane-spanning 4-domains, subfamily A	plasma membrane
	Seq ID 625 & 626	420610	AI683163	Hs.59348	distal-less homeo box 5	intracell
	Seq ID 627 & 628	425723	NM_014420	Hs.159311	clckapf (Xenopus laevis) homolog 4	secreted

Seq ID 629 & 630	456662	NM_002448	Hs.1494	msh (Drosophila) homeo box homolog 1 (fo	intracell
Seq ID 631 & 632	418281	U09550	Hs.1154	oviductal glycoprotein 1, 120kD (mucin 9	secreted
Seq ID 633 & 634	429903	AL134197	Hs.93597	cyclin-dependent kinase 5, regulatory su	Intracell

5

Table 76B:

Pkey: Unique Eos probeset identifier number

CAT number: Gene cluster number

Accession: Genbank accession numbers

10

Pkey	CAT Number	Accession
424399	238961_1	AI905687 AI905824 AI905837 AI905623 AA340069 R75793 W72837 BE074512 AI905633 W72838 BE092421 AI127172 BE186013 AW070916
		AI139456 AW176044 AW281950
428220	301384_1	AW207206 AW341473 AA448195 AI961341
450375	83327_1	AA009647 AA131254 AA374293 AW954405 HD4410 AW606284 AA151166 BE157467 BE157601 H04384 W48291 AW663674 H04021 H01532
		AA190993 H03231 H59605 H01642 AA852876 AA113758 AA626915 AA746952 AI161014 AA099554 R69067

15

Table 76C:

Pkey: Unique number corresponding to an Eos probeset

Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-495.

Strand: Indicates DNA strand from which exons were predicted.

NI_position: Indicates nucleotide positions of predicted exons.

20

25

Pkey	Ref	Strand	NI_position
402075	8117407	Plus	121907-122035,122804-122921,124019-124161,124455-124610,125672-126076
402230	9968312	Minus	29782-29832
403047	3540153	Minus	59793-59968
404029	7671252	Plus	108716-111112
404049	3888074	Minus	75765-78155
404682	9787231	Minus	40977-41160
404875	9801324	Plus	96588-96732,97722-97831
404877	1519284	Plus	1095-2107
404977	3738341	Minus	43081-43229
406932	7767812	Minus	123525-123713
406400	9256298	Plus	1553-1712,1878-2140,4252-4385,5922-6077

30

35

Table 77 provides Pkey; Seq ID No; Disease Indications; and Preferred Utility for sequences in Table 78. Seq ID No links the information in Table 77 to Table 78.

5	Seq ID No:	Sequence ID No for sequences in table		
	Pkey:	Unique Eos probaset identifier number		
15	Disease Indications:	Diseases designated for coverage as described in Table 1		
	Preferred Utility:	Preference of utility, based partly upon predicted localization (Ab is antibody; sm is small molecule target; CTL is vaccine target)		
	Seq ID No	Pkey	Disease Indications	Preferred Utility
10	Seq ID 1 & 2	425023	angiogenesis	Ab, sm, CTL, imaging
	Seq ID 3 & 4	424503	pancreas, prostate, angiogenesis, bladder, lung	Ab, sm, imaging
	Seq ID 5 & 6	429423	angiogenesis	Ab, sm
15	Seq ID 7 & 8	400289	angiogenesis, bladder, lung, cervical, ovarian, head & neck	Ab, sm, CTL, diagnostic
	Seq ID 9 & 10	419172	angiogenesis, renal	Ab, sm, CTL, imaging
	Seq ID 11 & 12	418007	angiogenesis, lung, bladder, fibrosis, head & neck, pancreas, stomach, colon, ovarian	Ab, sm, diagnostic
20	Seq ID 13 & 14	418007	angiogenesis, lung, bladder, fibrosis, head & neck, pancreas, stomach, colon, ovarian	Ab, sm, diagnostic
	Seq ID 15 & 16	407836	angiogenesis	CTL
	Seq ID 17 & 18	414577	angiogenesis	Ab, CTL, diagnostic
25	Seq ID 19 & 20	418738	angiogenesis, lung, ovarian, bladder & stomach, pancreas, uterine	Ab, sm, CTL, imaging
	Seq ID 21 & 22	428368	angiogenesis, head & neck, stomach	Ab, sm, diagnostic
	Seq ID 23 & 24	415138	angiogenesis, pancreas, stomach, lung, uterine	Ab, CTL, diagnostic
30	Seq ID 25 & 26	428276	angiogenesis, bladder, glioblastoma	Ab, sm, imaging
	Seq ID 27 & 28	418994	prostate, angiogenesis	Ab, CTL, imaging
	Seq ID 29 & 30	407975	angiogenesis, renal	Ab, CTL, diagnostic
35	Seq ID 31 & 32	429113	angiogenesis, bladder and stomach	sm, CTL
	Seq ID 33 & 34	418506	angiogenesis, ovarian, glioblastoma, uterine, lung, bladder, pancreas	Ab, sm, imaging
	Seq ID 35 & 36	423981	breast, colon, bladder, lung, fibrosis, pancreas, head and neck, ovarian	Ab, sm, diagnostic
40	Seq ID 37 & 38	414812	breast, bladder, lung, Fibrosis, pancreas, colon, head and neck, cervical, stomach, renal, ovarian	Ab, CTL, diagnostic
	Seq ID 39 & 40	417433	pancreas, breast, bladder, lung, head & neck, cervical, uterine, ovarian, stomach, renal	Ab, CTL, imaging
	Seq ID 41 & 42	417433	pancreas, breast, bladder, lung, head & neck, cervical, uterine, ovarian, stomach, renal	Ab, CTL, imaging
45	Seq ID 43 & 44	424399	breast, uterine, head & neck	Ab, CTL, diagnostic
	Seq ID 45 & 46	422867	breast, ovarian, prostate, pancreas, lung, colon, uterine	Ab, CTL, diagnostic
	Seq ID 47 & 48	428227	breast, lung, bladder, ovarian, head & neck, fibrosis, colon, stomach, cervical	Ab, CTL, diagnostic
50	Seq ID 49 & 50	444381	breast, colon, bladder, lung, pancreas, head & neck, ovarian, stomach, uterine, renal, angiogenesis	Ab, CTL, diagnostic
	Seq ID 51 & 52	439569	breast, androgen withdrawal prostate, prostate, bladder	Ab, CTL, diagnostic
	Seq ID 53 & 54	411558	pancreas, prostate, stomach, breast, uterine, cervical, ovarian	Ab, sm, imaging
55	Seq ID 55 & 56	400303	breast, ovarian, pros, stomach, uterine, bladder lung head & neck	Ab, sm, CTL, imaging
	Seq ID 57 & 58	411789	pancreas, lung, breast, stomach, head & neck, ovarian, uterine	Ab, CTL, diagnostic
	Seq ID 59 & 60	428688	breast, colon, lung, pancreas, stomach, head & neck, ovarian	Ab, sm, imaging
60	Seq ID 61 & 62	450098	breast, lung, stomach, uterine	CTL
	Seq ID 63 & 64	421552	breast, ovarian, pancreas, cervical, uterine, prostate, lung, stomach, head & neck	Ab, sm, CTL, diagnostic
	Seq ID 65 & 66	452747	breast, bladder, lung, head & neck, ovarian, stomach, uterine, pancreas	Ab, imaging
65	Seq ID 67 & 68	415539	breast, prostate, ovarian, cervical, uterine	Ab, sm, CTL, imaging
	Seq ID 69 & 70	418636	breast, pancreas, uterine	Ab, sm, CTL, imaging
	Seq ID 71 & 72	418636	breast, pancreas, uterine	Ab, sm, CTL, imaging
70	Seq ID 73 & 74	409079	breast, prostate	Ab, sm, CTL, imaging
	Seq ID 75 & 76	442082	breast, prostate, ovarian	Ab, imaging
	Seq ID 77 & 78	400297	breast, bladder, colon, prostate	Ab, sm, CTL, imaging
75	Seq ID 79 & 80	451398	breast, ovarian	CTL
	Seq ID 81 & 82	429220	breast, prostate, benign prostatic hyperplasia	Ab, CTL, imaging
	Seq ID 83 & 84	421624	breast	Ab, sm, CTL, imaging
80	Seq ID 85 & 86	423242	breast, renal, ovarian, prostate, colon	Ab, CTL, imaging
	Seq ID 87 & 88	423242	breast, renal, ovarian, prostate, colon	CTL
	Seq ID 89 & 90	423242	breast, renal, ovarian, prostate, colon	CTL
85	Seq ID 91 & 92	452190	breast, stomach, pancreas	CTL
	Seq ID 93 & 94	452190	breast, stomach, pancreas	CTL
	Seq ID 95 & 96	325372	breast	CTL
90	Seq ID 97 & 98	450375	breast, ovarian, head & neck, pancreas, lung, colon	Ab, sm, CTL, imaging
	Seq ID 99 & 100	426215	breast, lung, renal, colon, ovarian, uterine	Ab, CTL, diagnostic
	Seq ID 101 & 102	425247	breast, ovarian, lung, colon, pancreas, head & neck, stomach, uterine, cervical, bladder	Ab, sm, CTL, diagnostic
95	Seq ID 103 & 104	429353	breast, prostate	Ab, sm, CTL, imaging
	Seq ID 105 & 106	429353	breast, prostate	Ab, sm, CTL, imaging
	Seq ID 107 & 108	432201	breast, colon, lung chemo, ovarian, stomach, pancreas, uterine, cervical	Ab, sm, CTL, imaging
100	Seq ID 109 & 110	427585	breast, lung, head & neck, pancreas, stomach, colon, ovarian, cervical	CTL
	Seq ID 111 & 112	446163	breast, cervical, uterine	Ab, sm, imaging
	Seq ID 113 & 114	442117	breast, lung, bladder, pancreas, head & neck, stomach, ovarian, prostate	Ab, CTL, imaging
105	Seq ID 115 & 116	428179	breast	sm, CTL
	Seq ID 117 & 118	431211	colon, bladder, lung, pancreas, head & neck	Ab, sm, imaging
	Seq ID 119 & 120	447033	colon, prostate, fibrosis, breast, lung pancreas	Ab, CTL, diagnostic
110	Seq ID 121 & 122	447033	colon, prostate, fibrosis, breast, lung pancreas	Ab, CTL, diagnostic
	Seq ID 123 & 124	447033	colon, prostate, fibrosis, breast, lung pancreas	Ab, CTL, diagnostic
	Seq ID 125 & 126	115522	colon, lung, bladder, pancreas	CTL
115	Seq ID 127 & 128	452679	prostate, colon, pancreas, Taxol prostate	Ab, CTL, imaging
	Seq ID 129 & 130	446051	colon, breast	Ab, sm, CTL, imaging
	Seq ID 131 & 132	422048	colon, pancreas, prostate	diagnostic
120	Seq ID 133 & 134	410418	colon, bladder, lung, ovarian, pancreas, head & neck	Ab, sm, CTL, imaging
	Seq ID 135 & 136	446342	uterine, colon, prostate	Ab, sm, CTL, imaging
	Seq ID 137 & 138	422260	colon, ovarian mucinous	Ab, sm, CTL, diagnostic
125	Seq ID 139 & 140	409041	uterine, ovarian, lung, colon, stomach, head & neck, breast, pancreas	Ab, CTL, diagnostic
	Seq ID 141 & 142	409041	uterine, ovarian, lung, colon, stomach, head & neck, breast, pancreas	Ab, CTL, diagnostic
	Seq ID 143 & 144	111929	colon, breast, fibrosis	Ab, sm, imaging
130	Seq ID 145 & 146	111929	colon, breast, fibrosis	Ab, sm, imaging
	Seq ID 147 & 148	111929	colon, breast, fibrosis	Ab, sm, imaging

	Seq ID 149 & 150 104888	colon, stomach, uterine	Ab, imaging
	Seq ID 151 & 152 420159	bladder, stomach	Ab, sm, CTL, imaging
	Seq ID 153 & 154 422330	pancreas, colon, bladder	Ab, sm, CTL, imaging, diagnostic
5	Seq ID 155 & 156 452461	bladder, lung, head & neck, ovarian, glioblastoma, stomach, colon, cervical	CTL
	Seq ID 157 & 158 413324	bladder	Ab, sm, CTL, diagnostic
	Seq ID 159 & 160 412420	bladder, glioblastoma, lung, stomach	Ab, diagnostic
	Seq ID 161 & 162 416658	lung, ovarian, uterine, bladder	Ab, CTL, diagnostic
	Seq ID 163 & 164 407811	bladder, pancreas, stomach, uterine, lung	Ab, sm, diagnostic
10	Seq ID 165 402230	bladder	sm, CTL
	Seq ID 166 & 167 402230	bladder	sm, CTL
	Seq ID 168 & 169 432828	bladder	CTL
	Seq ID 170 & 171 425721	bladder	Ab, imaging
	Seq ID 172 & 173 420370	bladder	Ab, CTL, imaging
15	Seq ID 174 & 175 437852	bladder, lung	Ab, sm, CTL, imaging
	Seq ID 176 & 177 402075	bladder, lung, head & neck, cervical	diagnostic
	Seq ID 178 & 179 421110	bladder, pancreas, stomach, ovarian, lung	Ab, sm, diagnostic
	Seq ID 180 & 181 451668	bladder, ovarian, lung	Ab, sm, CTL, diagnostic
	Seq ID 182 & 183 451668	bladder, ovarian, lung	Ab, sm, CTL, diagnostic
20	Seq ID 184 & 185 451668	bladder, ovarian, lung	Ab, sm, CTL, diagnostic
	Seq ID 186 & 187 408243	bladder, stomach, head & neck, cervical	Ab, diagnostic
	Seq ID 188 & 189 422282	bladder, lung, head & neck	CTL, diagnostic
	Seq ID 190 & 191 425852	bladder, lung, head & neck	Ab, sm, CTL, imaging
	Seq ID 192 & 193 439738	bladder, lung, cervical	Ab, sm, CTL, imaging
25	Seq ID 194 & 195 404875	bladder	sm, CTL
	Seq ID 196 & 197 425883	bladder, pancreas	Ab, CTL, imaging
	Seq ID 198 & 199 404977	bladder, ovarian	Ab, sm, CTL, diagnostic
	Seq ID 200 & 201 420876	pancreas, bladder	Ab, sm, CTL, imaging
	Seq ID 202 & 203 427747	bladder, lung, ovarian, stomach	sm, CTL
30	Seq ID 204 & 205 420281	lung, bladder, ovarian, pancreas	Ab, sm, imaging
	Seq ID 206 & 207 446573	bladder	sm, CTL
	Seq ID 208 437553	bladder	Ab, CTL, imaging
	Seq ID 209 & 210 437553	bladder	Ab, CTL, imaging
	Seq ID 211 & 212 437553	bladder	Ab, CTL, imaging
35	Seq ID 213 & 214 426800	bladder, prostate	Ab, CTL, imaging
	Seq ID 215 & 216 426800	bladder, prostate	Ab, CTL, imaging
	Seq ID 217 & 218 426800	bladder, prostate	Ab, CTL, imaging
	Seq ID 219 & 220 405932	bladder, lung, head & neck, cervical	Ab, CTL, imaging
	Seq ID 221 & 222 405932	bladder, lung, head & neck, cervical	sm
40	Seq ID 223 & 224 424008	bladder, head & neck, stomach, cervical	sm
	Seq ID 225 & 226 424008	bladder, head & neck, stomach, cervical	Ab, sm, CTL, imaging
	Seq ID 227 & 228 424008	bladder, head & neck, stomach, cervical	Ab, sm, CTL, imaging
	Seq ID 229 & 230 444342	hepatitis C, lung, fibrosis, bladder	Ab, CTL, diagnostic
	Seq ID 231 & 232 421379	breast, pancreas, head & neck, lung, stomach, bladder, cervical, colon	Ab, CTL, diagnostic
45	Seq ID 233 & 234 417079	bladder, lung, head & neck, cervical	Ab, diagnostic
	Seq ID 235 & 236 412886	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 237 & 238 412886	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 239 & 240 412886	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 241 & 242 447072	glioblastoma, pancreas	sm, CTL
50	Seq ID 243 & 244 419723	glioblastoma	Ab, CTL, diagnostic
	Seq ID 245 & 246 419723	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 247 & 248 430890	glioblastoma, lung, cervical, bladder	Ab, CTL, imaging, diagnostic
	Seq ID 249 & 250 456759	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 251 429466	glioblastoma, uterine	
55	Seq ID 252 429466	glioblastoma, uterine	
	Seq ID 253 & 254 419721	glioblastoma, lung	Ab, sm, CTL, imaging
	Seq ID 255 & 256 407034	glioblastoma	Ab, CTL, diagnostic
	Seq ID 257 & 258 413472	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 259 & 260 438380	glioblastoma	Ab, CTL, diagnostic, imaging
60	Seq ID 261 & 262 426271	glioblastoma	Ab, CTL, diagnostic
	Seq ID 263 & 264 419704	glioblastoma	sm, CTL
	Seq ID 265 & 266 444471	glioblastoma, lung, colon	Ab, sm, CTL, imaging
	Seq ID 267 & 268 405935	glioblastoma	Ab, CTL, diagnostic
	Seq ID 269 & 270 413063	glioblastoma, ovarian, bladder, lung	Ab, CTL, diagnostic
65	Seq ID 271 & 272 433800	glioblastoma, lung	Ab, CTL, imaging
	Seq ID 273 458435	glioblastoma	
	Seq ID 274 458435	glioblastoma	
	Seq ID 275 & 276 424343	glioblastoma, ovarian, uterine	Ab, sm, CTL, imaging
	Seq ID 277 & 278 424988	glioblastoma	Ab, CTL, diagnostic
70	Seq ID 279 & 280 412709	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 281 & 282 435615	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 283 & 284 404049	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 285 & 286 418932	glioblastoma	Ab, sm, CTL, imaging
	Seq ID 287 & 288 404029	glioblastoma	sm, CTL
75	Seq ID 289 & 290 436480	glioblastoma	Ab, CTL, diagnostic
	Seq ID 291 & 292 452401	bladder, breast, pancreas, head & neck, stomach, lung, arthritis, renal	Ab, CTL, diagnostic
	Seq ID 293 & 294 452401	bladder, breast, pancreas, head & neck, stomach, lung, arthritis, renal	Ab, sm, imaging
	Seq ID 295 & 296 438895	breast, renal, ovarian, glioblastoma	Ab, sm, CTL, imaging
	Seq ID 297 & 298 421471	renal	Ab, sm, CTL, imaging
80	Seq ID 299 & 300 428296	renal	Ab, sm, CTL, imaging
	Seq ID 301 & 302 423508	renal, colon	Ab, CTL, imaging
	Seq ID 303 & 304 450001	renal, lung	Ab, sm, CTL, imaging
	Seq ID 305 & 306 410407	renal, lung, colon, stomach, ovarian, uterine	Ab, sm, CTL, imaging
	Seq ID 307 & 308 453496	renal, prostate	Ab, sm, CTL, imaging

	Seq ID 309 & 310 420737	renal	Ab,sm, CTL, imaging
	Seq ID 311 & 312 309931	lung	Ab,sm, CTL, imaging
	Seq ID 313 & 314 412719	lung, head & neck, bladder, glioblastoma, cervical	CTL
5	Seq ID 315 & 316 417034	lung, head & neck, a couple cervical	Ab, CTL, diagnostic
	Seq ID 317 & 318 430486	lung, bladder, head & neck, cervical	Ab,sm, imaging
	Seq ID 319 & 320 413753	lung, bladder, head & neck, pancreas, cervical, colon	CTL, diagnostic
	Seq ID 321 & 322 425650	lung, head & neck, cervical, bladder	Ab, imaging
	Seq ID 323 & 324 423673	bladder, lung, head & neck, ovarian, pancreas, colon, stomach, uterine, cervical	Ab, CTL, diagnostic
10	Seq ID 325 & 326 418683	lung, bladder, head & neck, cervical	Ab, imaging
	Seq ID 327 & 328 418683	lung, bladder, head & neck, cervical	Ab, imaging
	Seq ID 329 & 330 428610	lung	CTL, diagnostic
	Seq ID 331 & 332 406690	lung, head & neck, pancreas, stomach, bladder, colon, cervical	Ab, imaging
	Seq ID 333 & 334 431846	lung, bladder, head & neck, uterine, cervical, stomach, ovarian	Ab, imaging
15	Seq ID 335 & 336 422158	head & neck, bladder, lung, cervical, stomach	diagnostic
	Seq ID 337 & 338 431958	lung, bladder, cervical, head & neck, ovarian, colon, prostate, pancreas, breast	Ab, CTL, imaging
	Seq ID 339 & 340 437044	head & neck, cervical, lung, bladder, breast, prostate, ovarian, stomach	Ab, imaging
	Seq ID 341 & 342 428484	lung, glioblastoma, bladder, head & neck, colon, cervical	Ab, sm, imaging
	Seq ID 343 & 344 429211	lung, bladder, head & neck, cervical, stomach	Ab, imaging
20	Seq ID 345 & 346 417389	ovarian, lung, bladder, uterine, cervical, pancreas, stomach	Ab, diagnostic
	Seq ID 347 & 348 431009	lung, bladder, head & neck	Ab, sm, imaging
	Seq ID 349 & 350 417542	lung	CTL, diagnostic
	Seq ID 351 & 352 449230	lung, cervical, head & neck, bladder, ovarian, colon	Ab, imaging
	Seq ID 353 & 354 410555	lung	Ab, sm, imaging
25	Seq ID 355 & 356 410555	lung	Ab, sm, imaging
	Seq ID 357 & 358 424887	head & neck, pancreas, lung, uterine, cervical, colon, stomach	Ab, sm, diagnostic
	Seq ID 359 & 360 418462	lung, bladder	Ab, imaging
	Seq ID 361 & 362 410274	lung, renal	diagnostic
	Seq ID 363 & 364 439606	lung, bladder, head & neck, cervical	Ab,sm, imaging
30	Seq ID 365 & 366 404877	lung, bladder	CTL
	Seq ID 367 & 368 444781	lung, bladder, head & neck, cervical	Ab, imaging
	Seq ID 369 & 370 109424	bladder, lung	Ab, sm, CTL, imaging
	Seq ID 371 & 372 415817	lung, glioblastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
	Seq ID 373 & 374 415817	lung, glioblastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
35	Seq ID 375 & 376 416817	lung, glioblastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
	Seq ID 377 & 378 415817	lung, glioblastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
	Seq ID 379 & 380 415817	lung, glioblastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
	Seq ID 381 & 382 415817	lung, glioblastoma, head & neck, cervical, fibrosis	Ab,sm, CTL, imaging
	Seq ID 383 & 384 421817	lung, cervical, head & neck & bladder	Ab,sm, CTL, imaging
40	Seq ID 385 & 386 418678	lung, bladder, stomach, ovarian, pancreas & cervical	Ab,sm, CTL, imaging
	Seq ID 387 & 388 418678	lung, bladder, stomach, ovarian, pancreas & cervical	CTL
	Seq ID 389 & 390 409420	lung, head & neck, pancreas, stomach, cervical, bladder	CTL, diagnostic
	Seq ID 391 & 392 332180	lung	Ab, sm, imaging
	Seq ID 393 & 394 408790	lung	Ab,sm, CTL, imaging
45	Seq ID 395 & 396 408790	lung	Ab,sm, CTL, imaging
	Seq ID 397 & 398 439223	lung, head & neck, cervical, bladder & colon	Ab, CTL, imaging
	Seq ID 399 & 400 409757	pancreas, stomach, lung, bladder, stomach	Ab, CTL, diagnostic
	Seq ID 401 & 402 428969	lung, cervical	Ab, CTL, diagnostic
	Seq ID 403 & 404 428969	lung, cervical	Ab, CTL, diagnostic
50	Seq ID 405 & 406 428969	lung, cervical	Ab, CTL, diagnostic
	Seq ID 407 & 408 428969	lung, cervical	Ab, CTL, diagnostic
	Seq ID 409 & 410 450701	lung, head & neck, pancreas, breast, stomach, ovarian, colon	Ab, CTL, diagnostic
	Seq ID 411 & 412 450701	lung, head & neck, pancreas, breast, stomach, ovarian, colon	Ab, CTL, diagnostic
	Seq ID 413 & 414 414774	lung, bladder, head & neck, pancreas, stomach, ovarian	Ab, CTL, diagnostic
55	Seq ID 415 & 416 407944	lung, head & neck	Ab, sm, diagnostic
	Seq ID 417 & 418 407944	lung, head & neck	Ab, sm, imaging
	Seq ID 419 & 420 428486	pancreas	Ab, sm, imaging
	Seq ID 421 & 422 457489	pancreas, prostate, lung	Ab, CTL, diagnostic
	Seq ID 423 & 424 432874	pancreas, stomach	Ab, CTL, diagnostic
60	Seq ID 425 & 426 445891	stomach, pancreas, ovarian	Ab, CTL, imaging
	Seq ID 427 & 428 445891	stomach, pancreas, ovarian	Ab, CTL, imaging
	Seq ID 429 & 430 404682	pancreas	Ab, CTL, diagnostic
	Seq ID 431 & 432 429547	pancreas, head & neck, lung, ovarian	Ab, CTL, diagnostic
	Seq ID 433 & 434 425821	stomach, pancreas	Ab,sm, CTL, imaging
65	Seq ID 435 & 436 407242	pancreas, colon, bladder, head & neck, stomach, lung, ovarian, cervical	Ab,sm, CTL, imaging
	Seq ID 437 & 438 407242	pancreas, colon, bladder, head & neck, stomach, lung, ovarian, cervical	Ab,sm, CTL, imaging
	Seq ID 439 & 440 407242	pancreas, colon, bladder, head & neck, stomach, lung, ovarian, cervical	Ab,sm, CTL, imaging
	Seq ID 441 & 442 432596	pancreas, breast	CTL
	Seq ID 443 & 444 444006	pancreas, colon, lung, ovarian & cervical	Ab, CTL, imaging
70	Seq ID 445 & 446 423686	pancreas, uterine, colon	CTL
	Seq ID 447 & 448 428382	pancreas	Ab, CTL, diagnostic
	Seq ID 449 & 450 429587	pancreas, colon, stomach, lung	Ab,sm, CTL, imaging
	Seq ID 451 & 452 448030	pancreas, renal and stomach	Ab,sm, imaging
	Seq ID 453 & 454 422109	pancreas, lung, colon	Ab, CTL, diagnostic
75	Seq ID 455 & 456 419235	pancreas, fibrosis, head & neck & lung	Ab, CTL, imaging
	Seq ID 457 & 458 449048	pancreas, ovarian, uterine, glioblastoma, head & neck & lung	Ab, CTL, imaging
	Seq ID 459 & 460 427333	pancreas, colon	Ab, sm, imaging
	Seq ID 461 & 462 417931	ovarian, pancreas, stomach, colon, uterine, prostate	Ab, diagnostic
	Seq ID 463 & 464 419216	pancreas, lung, stomach, cervical, prostate, head & neck	Ab, CTL, diagnostic
80	Seq ID 465 & 466 431629	pancreas, uterine, cervical, stomach	Ab, CTL, diagnostic
	Seq ID 467 & 468 413554	pancreas, glioblastoma	Ab, CTL, diagnostic
	Seq ID 469 & 470 452194	stomach, pancreas, renal, colon	Ab,sm, CTL, imaging
	Seq ID 471 & 472 452194	stomach, pancreas, renal, colon	Ab,sm, CTL, imaging
	Seq ID 473 & 474 426322	pancreas, bladder, stomach	Ab, diagnostic

	Seq ID 475 & 476 429010	pancreas	sm, CTL
	Seq ID 477 & 478 431462	pancreas, lung, glioblastoma	Ab, diagnostic
	Seq ID 479 & 480 448243	ovarian, uterine, lung, stomach, head & neck, glioblastoma, pancreas	Ab, sm, imaging
5	Seq ID 481 & 482 428427	ovarian, lung, head & neck, cervical, colon, uterine, stomach	sm, CTL
	Seq ID 483 & 484 428187	ovarian, uterine, colon, stomach	Ab, sm, CTL, imaging
	Seq ID 485 & 486 428579	ovarian, Ewing, uterine, pancreas	Ab, sm, CTL, imaging
	Seq ID 487 & 488 428579	ovarian, Ewing, uterine, pancreas	Ab, sm, CTL, imaging
	Seq ID 489 & 490 428579	ovarian, Ewing, uterine, pancreas	Ab, sm, CTL, imaging
10	Seq ID 491 & 492 428579	ovarian, Ewing, uterine, pancreas	Ab, sm, CTL, imaging
	Seq ID 493 & 494 445537	ovarian, bladder, uterine, breast, lung, head & neck, renal, fibrosis, pancreas, cervical	Ab, CTL, diagnostic
	Seq ID 495 & 496 422278	ovarian, head & neck, bladder, cervical, lung	Ab, sm, imaging
	Seq ID 497 & 498 424620	ovarian	Ab, CTL, diagnostic
	Seq ID 499 & 500 406400	ovarian, uterine	Ab, CTL, diagnostic
15	Seq ID 501 & 502 431130	ovarian, uterine	Ab, CTL, diagnostic
	Seq ID 503 & 504 420440	ovarian, uterine, cervical	Ab, CTL, diagnostic
	Seq ID 505 & 506 428450	ovarian, cervical, pancreas, lung	sm
	Seq ID 507 & 508 446619	ovarian, fibrosis, pancreas, head & neck, lung, colon	Ab, diagnostic
	Seq ID 509 & 510 436982	ovarian, fibrosis	Ab, CTL, diagnostic
20	Seq ID 511 & 512 453392	ovarian, lung, glioblastoma	CTL
	Seq ID 513 431989	ovarian	
	Seq ID 514 439820	ovarian, uterine, cervical, breast, prostate	
	Seq ID 515 & 516 409178	ovarian, breast	Ab, CTL, diagnostic
	Seq ID 517 & 518 426514	ovarian, colon, bladder, lung, cervical	Ab, CTL, diagnostic
25	Seq ID 519 & 520 421478	ovarian, bladder	Ab, CTL, imaging
	Seq ID 521 & 522 421478	ovarian, bladder	Ab, CTL, imaging
	Seq ID 523 & 524 425776	ovarian, uterine, lung	Ab, sm, CTL, imaging
	Seq ID 525 & 526 425776	ovarian, uterine, lung	Ab, sm, CTL, imaging
	Seq ID 527 & 528 452097	ovarian	Ab, sm, diagnostic
30	Seq ID 529 & 530 416530	ovarian, uterine	Ab, CTL, diagnostic
	Seq ID 531 & 532 431515	ovarian, prostate, lung	Ab, sm, CTL, imaging
	Seq ID 533 & 534 419452	ovarian, prostate, lung, breast, uterine	Ab, sm, CTL, imaging
	Seq ID 535 & 536 412078	ovarian	CTL
	Seq ID 537 & 538 412078	ovarian	CTL
35	Seq ID 539 & 540 411773	ovarian	Ab, sm, CTL, diagnostic
	Seq ID 541 & 542 407792	ovarian, uterine, cervical, pancreas	Ab, CTL, diagnostic
	Seq ID 543 & 544 431616	prostate, pancreas, colon	Ab, sm, CTL, imaging
	Seq ID 545 & 546 452792	prostate, uterine, breast	Ab, CTL, imaging
	Seq ID 547 & 548 400294	prostate, taxol prostate	Ab, sm, CTL, imaging
40	Seq ID 549 & 550 432653	prostate, lung	Ab, sm, CTL, imaging
	Seq ID 551 & 552 432653	prostate, lung	Ab, sm, CTL, imaging
	Seq ID 553 & 554 432653	prostate, lung	Ab, sm, CTL, imaging
	Seq ID 555 & 556 432653	prostate, lung	Ab, sm, CTL, imaging
	Seq ID 557 & 558 400290	prostate, colon	Ab, sm, CTL, imaging
45	Seq ID 559 & 560 410001	ovarian, prostate, uterine, cervical, lung	Ab, CTL, diagnostic
	Seq ID 561 & 562 418396	prostate	Ab, sm, CTL, imaging
	Seq ID 563 & 564 451027	prostate, uterine, glioblastoma	Ab, sm, CTL, imaging
	Seq ID 565 & 566 446057	prostate	Ab, sm, CTL, imaging
	Seq ID 567 & 568 433486	prostate	Ab, CTL, diagnostic
	Seq ID 569 & 570 453370	prostate	Ab, sm, CTL, imaging
50	Seq ID 571 & 572 453370	prostate	Ab, sm, CTL, imaging
	Seq ID 573 & 574 414569	prostate	Ab, sm, CTL, imaging
	Seq ID 575 & 576 413435	prostate, glioblastoma, pancreas	Ab, sm, diagnostic
	Seq ID 577 & 578 426501	prostate, breast, glioblastoma, lung	Ab, CTL, imaging
55	Seq ID 579 & 580 448999	prostate, glioblastoma	Ab, sm, CTL, imaging
	Seq ID 581 & 582 408369	prostate, lung, fibrosis, uterine, glioblastoma, cervical, ovarian	Ab, sm, CTL, imaging
	Seq ID 583 & 584 412628	prostate	Ab, CTL, diagnostic
	Seq ID 585 & 586 403047	prostate, bladder, colon	Ab, sm, CTL, imaging
	Seq ID 587 & 588 403047	prostate, bladder, colon	Ab, sm, CTL, imaging
60	Seq ID 589 & 590 408430	prostate	Ab, sm, CTL, imaging
	Seq ID 591 & 592 445413	prostate, colon, uterine, ovarian, lung, pancreas	diagnostic
	Seq ID 593 & 594 451982	prostate, bladder	CTL
	Seq ID 595 & 596 427958	prostate, glioblastoma	Ab, sm, CTL, imaging
	Seq ID 597 & 598 421887	prostate	Ab, CTL, imaging
65	Seq ID 599 & 600 425071	prostate, colon, stomach, uterine, cervical, head & neck, pancreas	Ab, diagnostic
	Seq ID 601 & 602 432101	prostate, pancreas	Ab, sm, imaging
	Seq ID 603 & 604 407786	prostate, colon, uterine, stomach, inflammatory bowel disease, ovarian	Ab, sm, imaging
	Seq ID 605 & 606 416838	Prostate, Ewing, glioblastoma	Ab, CTL, diagnostic
	Seq ID 607 & 608 416539	ZD1839 resistant cancers, head & neck	Ab, sm, CTL, imaging
70	Seq ID 609 & 610 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
	Seq ID 611 & 612 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
	Seq ID 613 & 614 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
	Seq ID 615 & 616 131083	androgen withdrawal prostate	Ab, sm, CTL, imaging
	Seq ID 617 & 618 422424	bladder, pancreas, prostate, angiogenesis, colon, stomach, lung	Ab, CTL, diagnostic
75	Seq ID 619 & 620 426970	stomach, pancreas, colon	Ab, sm, imaging
	Seq ID 621 & 622 428330	uterine, ovarian, fibrosis, prostate, pancreas, lung, bladder, head & neck	Ab, sm, CTL, diagnostic
	Seq ID 623 & 624 439018	uterine, stomach, prostate	Ab, sm, CTL, imaging
	Seq ID 625 & 626 420810	uterine, ovarian endometrioid, lung	CTL
	Seq ID 627 & 628 425723	ovarian endometrioid, uterine, colon	Ab, CTL, diagnostic
80	Seq ID 629 & 630 458862	uterine, ovarian	CTL
	Seq ID 631 & 632 418281	uterine, ovarian	Ab, CTL, diagnostic
	Seq ID 633 & 634 429903	lung	sm

Table 78

Seq ID NO: 1 DNA sequence
Nucleic Acid Accession #: NM_001400
Coding sequence: 251..1399

	1	11	21	31	41	51	
5	TCTAAAGGTC	GGGGGCGAGCA	GCAAGATGCG	AAGCGAGCCG	TACAGATCCC	GGGCTCTCCG	60
10	AAAGCAACTT	CGCCCTGCTT	GAGCGAGGCT	GCGGTTCGCG	AGGCCCTCTC	CAGCCAAAGGA	120
	AAAGCTACAC	AAAGAGCCCTG	GATCACTCAT	CGAACCACCC	CTGAAGCCAG	TGAAGGCTCT	180
	CTGCGCTGCG	CCTCTAGCGT	TCGTCTGGAG	TAGCGCCACC	CGGCTTCCT	GGGACACAG	240
	GGTTGCGACC	ATGGGGCCCA	CCAGCGTCCC	GCTGGTCAAG	GCCACCCCA	GCTCGGTCTC	300
15	TGACTACGTC	AACTATGATA	TCATCGTCCG	GCATTACAAC	TACACCGGAA	AGCTGAATAT	360
	CAGCGCGAGC	AAGGAGAACA	GCATTAAACT	GACCTCGGTG	GTGTTTCATT	TCATCTGCTG	420
	CTTTATCATC	TGGGAGAACA	TCTTTGTCTT	GCTGACCATT	TGGAAAACA	AGAAATTCCT	480
	CCGACCCATG	TACTATTCTA	TGGCAATCT	GGCCCTCTCA	GACCTGTGG	CAGGAGTAGC	540
	CTACACAGCT	AACCTGCTCT	TGTCTGGGGC	CACACCTAC	AAGCTCACTC	CCGCCAGTG	600
20	GTTCCTGCGG	GAAGGGAGTA	TGTTTGTGGC	CCTGTGAGCC	TCCGTGTTCA	GTCTCCTCCT	660
	CATCGCCATT	GAGGCTATA	TCACAATGCT	GAAATGAAA	CTCCACAACG	GGAGCAATAA	720
	CTTCGCGCTC	TTCCCTGCTA	TCAGCGCTCG	CTGGGTGATC	TCCCTCATCC	TGGGTGGCCT	780
	GCCTATCATG	GGCTGGAATC	GCATCAGTGC	GCTGTCCAGC	TGCTCCACCG	TGCTGCCGCT	840
	CTACCAACAG	CATCATATCC	TCTCTGCAC	CACGGTCTTC	ACTCTGCTTC	TGCTCTCCAT	900
25	CGTCATTCTG	TACTGCAGAA	TCTACTCCTT	GGTCAGGACT	CGAGCCGCC	GCCTGACGTT	960
	CCGCAAGAAC	ATTTCACAGG	CCAGCCGCG	CTCTGAGAAG	TGCTGGGCGC	TGCTCAAGAC	1020
	CGTAATTATC	GGCTTGAGCG	TCTTCATGCG	CTGCTGGGCA	CGGCTCTTCA	TCTGCTCCT	1080
	CGTGGATGCG	GGCTGCAAGG	TGAAGACCTG	TGACATCCTC	TTGAGAGCGG	AGTACTTCCT	1140
	GGTGTGAGCT	GTGCTCACT	CCGGCAACAA	CCCATCATTT	TACACTCTGA	CCACCAAGGA	1200
30	GATCGCTCGG	GCCTTCATCC	GGATCATGTC	CTGCTGCCAG	TGCCCCAGCG	GAGACTCTGC	1260
	TGGCAATTC	AAGCGACCCA	TCATCGCCCG	CATGGAATTC	AGCCGCGAGCA	AATCGGACAA	1320
	TTCTCTCCAC	CCCGAGAAAG	ACGAAGGGGA	CAACCCAGAG	ACCATTATGT	CTCTGGGAAA	1380
	CGTCAACTCT	TCTCTCTAGA	ACTGGAAGCT	GTCACCCAC	OGGAAGCCCT	CTTTACTTGG	1440
	TGCTCGGCGA	CCCGAGTGT	TGGAAAAAAA	TCTCTGGGCT	TGAGTCTCTG	CCAGGGAGGA	1500
35	GCTGCTGCAA	GCCAGAGGGG	GGAAGGGGGA	GAATACGAAC	AGCCTGGTGG	TCTCGGGTGT	1560
	TGGTGGGTAG	AGTTAGTTCC	TGTGAACAT	GCACCTGGAA	GGGTGGAGAT	CAGGTCCCGG	1620
	CTGGGAATAT	ATATTCTACC	CCCTCGAGC	TTTGATTTTG	CACCTAGCCA	AGGTCTAGC	1680
	ATTGTCAAGC	TTCTAAAGGG	TTCAATTGGC	CCCTCTCAA	AGACTAATGT	CCCATGTGA	1740
	AAGCTCTCT	TGCTCTGGAG	CTTTGAGGAG	ATGTTTCTCT	TCACTTTAGT	TTCAAAACCA	1800
40	AGTGAGTGTG	TGCACCTCTG	CTTCTTTAGG	GATGCCCTGT	ACATCCCA	CCCCACCTC	1860
	CTTCTCCTTC	ATACCCCTCC	TCACGTTCT	TTTACTTTAT	ACTTTAACA	CCTGAGAGTT	1920
	ATCAGAGCTG	GGGTTGTGGA	ATGATCGATC	ATCTATAGCA	AATAGGCTAT	GTGAGTACG	1980
	TAGGCTGTGG	GAAGATGAAG	ATGGTTTGA	GGTGTAAAC	AATGTCTTC	GCTGAGGCCA	2040
	AAGTTTCAT	GTAGCGGGA	TCCGTTTCT	GGAATTTGGT	TGAAGTCAT	TTGATTCTCT	2100
45	TAAAAACAT	CTTTTCAATG	AAATGRTTA	CCATTTCATA	TCCATTGAAG	CCGAATCTG	2160
	CATAGGAGAG	CCACATTTAT	CTAAATGATA	TTAGCCAGGA	TCCCTGGTGT	CCTAGGAGAA	2220
	ACAGACACAG	AAAACAAAGT	GAAAACCGAA	TGGATTAACT	TTTGCAAAAC	AAGGGAGATT	2280
	TCTTAGCAAA	TGAGCTTAAC	AAATATGACA	TCCGCTCTTC	CCACTTTTGT	TGATGTTTAT	2340
	TTCAAGATCT	TGTGTGATT	ATTCAAGCA	ACAACATGTT	GTAATTTGTT	GTGTTAAAG	2400
50	TACTTTTCTT	GATTTTGTAA	TGTATTGTT	TCAGGAAGAA	GTCATTTTAT	GGATTTTCTT	2460
	ATCCCGTGT	AACCTTTCTA	GAATCCACCC	TCTTGTGCC	TTAGCATTA	CTTTACTGCG	2520
	TAGGGAACGC	CAGACTTTT	AAGTCCAGCT	ATTCAATTGA	TAGTAATTGA	AGATATGTAT	2580
	AAATATTACA	AAGATAAAAA	ATATATTACT	GTCTCTTAG	TATGGTTTTC	AGTGCATTA	2640
	AACCGAGAGA	TGCTCTGTTT	TTTTAAAAAG	AATAGTATTT	AATAGGTTTC	TGACTTTTGT	2700
55	GGATCATTTT	GCACATAGCT	TTATCAACTT	TAAACATTA	ATAAATGAT	TTTTTTAAAG	

Seq ID NO: 2 Protein sequence
Protein Accession #: NP_001391

	1	11	21	31	41	51	
60	MGPTSVPLVK	AHRGSVSDYV	NYDIIVRHYN	YTGKLNISAD	KENSIKLTSTV	VFILICCFII	60
	LENIFVLLTI	WKTQKQFHRFM	YYFIGNLALS	DLLAGVAYTA	MLLSGATFY	KLTPAQWFLR	120
	EGSMFVALSA	SVFBLALAI	ERYITMLKMK	LHNGSNMFL	PLLISACNVI	SLILGGLPIM	180
65	QNNCISALSS	CSFVLPVHK	HYILFCTTVP	TLLLSIVIL	YCRIYSLVRT	RSRLTFRKN	240
	ISKASRSSEK	SLALLKTVII	VLSVFIACMA	PLFILLLLDV	GCKVKTCIDIL	FRABYFLVLA	300
	VLSGNTNPII	YTLNKEMRR	AFIRIMSCCK	CPSGDSAGKF	KRPILAGMEF	SRKSDNSSE	360
	PQDEDEDNPE	TIMSSGNVNS	SS				

Seq ID NO: 3 DNA sequence
Nucleic Acid Accession #: NM_002205.1
Coding sequence: 1..3149

	1	11	21	31	41	51	
75	ATGGGGAGCC	GGACGCCAGA	GTCCCTCTC	CACGCGGTGC	AGCTGCGCTG	GGGCCCCCGG	60
	CGCCGACCCC	CGCTSTTGC	GCTGCTGTTG	CTGCTSTTGC	CGCCGCCACC	CAGGCTGGG	120
	GGCTCTCACT	TAGACCGCGA	GGCCCGAGCA	GTACTCTCGG	GGCCCGCCGG	CTCCTCTCTC	180
	GGATTCTCAG	TGGATCTTTA	CCGGCCGGGA	ACAGACGGGG	TCAGTGTGCT	GGTGGGAGCA	240
80	CCCAAGGCTA	ATACCGAGCA	GCCAGGAGTG	CTGCAGGGTG	GTGCTGTCTA	CCTCTGTCTC	300
	TGGGGTGCCA	GCCCCACACA	GTGCACCCCC	ATTGAATTTG	ACAGCAAGGG	CTCTCGGCTC	360
	CTGGAGTCTC	CACCTGTCCAG	CTCAGAGGGA	GAGGAGCCTG	TGGAGTACAA	GTCTTGCAG	420
	TGGTTCGGGG	CAACAGTTCC	AGCCCATGGC	TCTTCCATCT	TGGCATGCGC	TCCACTGTAC	480
	AGCTGGCGCA	CAGAGAAGGA	GCCACTGAGC	GACCCGCTGG	GCACCTGCTA	CCTCTCCACA	540
	GATAACTTCA	CCGAATTCTT	GGAGATGCA	CCCTGCGGCT	CAGATTTCAG	CTGGCAGCA	600

5
10
15
20
25
30
35
40
45

```

GGACAGGGTT ACTGCCAAGG AGGCTTCAGT GCCGAGTTCA CCAAGACTGG CCGTGTGGTT 660
TTAGGTGGAC CAGGAAGCTA TTCTGGCAA GGCAGATCC TGTCTGCCAC TCAGGAGCAG 720
ATTGCAGAAAT CTTATTACCC CGAGTACCTG ATCAACCTGG TTCAGGGGCA GCTGCAGACT 780
CGCCAGGCCA GTTCCATCTA TGATGACAGC TACCTAGGAT ACTCTGTGGC TGTGTGGTGA 840
TTCAGTGGTG ATGACACAGA AGACTTTGTT GCTGGTGTGC CCAAGGGGAA CCTCACTTAC 900
GGCTATGTCA CCATCCTTAA TGGCTCAGAC ATTOGATCCC TCTACAACTT CTCAGGGGAA 960
CAGATGGCCT CTTACTTTGG CTATGCACTG GCCGCCACAG ACGTCAATGG GGAAGGGCTG 1020
GATGACTTGC TGGTGGGGGC ACCCTTGCCT ATGGATCGGA CCCCTGACGG GCGGCTCAG 1080
GAGGTGGGCA GGGTCTACGT CTACCTGCAG CACCCAGCCG GCATAGAGCC CAGGCCACC 1140
CTTACCTTCA CTGCGCATGA TGAGTTTGGC CGATTGGCA GCTCCTTGAC CCCCTGGGG 1200
GACCTGGACC AGGATGGCTA CAATGATGTG GCCATCGGG CTCCTTTGG TGGGAGACC 1260
CAGCAGGGAG TAGTGTTTGT ATTCTCTGG GGCAGGAG GGTGGGCTC TAAGCTTTC 1320
CAGGTCTTGC AGCCCTGTG GGCAGCCAGC CACACCCAG ACTTCTTGG CTCTGCCCT 1380
CGAGGAGGCC GAGACCTGGA TGCCAAATGA TATCTGATC TGATTGTGGG GTCCTTTGGT 1440
GTGGACAAGG CTGAGGCTGA CAGGGGCGCG CCATCGTGT CCGCTAGTGC CTCCTCACC 1500
ATCTTCCCGG CCATGTTCAA CCCAGAGGAG CGGAGCTGCA GCTTAGAGGG GAACCTGTG 1560
GCCTGCATCA ACCTTAGCTT CTGCCTCAAT GCTTCTGGA AACACGTTG TGACTCCATT 1620
GGTTTACAG TGGAACITCA GCTGGACTGG CAGAAGCAGA AGGGAGGGGT ACGGCGGGCA 1680
CTGTTCCTGG CCTCCAGCA GGCACCCCTG ACCCAGACCC TGCTCATCCA GAATGGGGCT 1740
CGAGAGGATT CGACAGAGAT GAAGATCTAC CTCAGGAACG AGTCAGAAAT TOGAGACAAA 1800
CTCTGCGCGA TTCCATCGC TCTCAACTTC TCCTTGGACC CCCAGCCCG AGTGGACAGC 1860
CACGGCCTCA GGCAGCCCTT ACATTATCAG AGCAGAGACC GGATAGAGGA CAGGGCTCAG 1920
ATCTTGTCTG ACTGTGGAGA AGACAACATC TGTGTGCTTG ACTTGTGGT GGAAGTGT 1980
GGGGAGCAGA ACCATGTGTA CCGGGGTGAC AAGAAATGCC TGAACTCAC TTTCCATGCC 2040
CAGAATGTGG GTGAGGGTGG CGCCTATGAG GCTGAGCTTC GGTCAACGCG CCTCCAGAG 2100
GCTGAGTACT CAGGACTCGT CAGACACCCA GGGAACTTCT CCAGCCTGAG CTGTGACTAC 2160
TTTGCGGTGA ACCAGAGCGG CCTGCTGCTG TGTGACCTGG GCAACCCCAT GAAGGCAGGA 2220
GCCAGTCTGT GGGGTGGGCT TCGGTTTACA GTCCCTCATC TCGGGGACAC TAAGAAAACC 2280
ATCCAGTTTG ACTTCCAGAT CCTCAGCAAG AATCTCAACA ACTCGCAAG CGAGTGGTT 2340
TCTTTTCGGG TCTCGTGGG GCTCAGGCC CAGGTCAACC TGAACGGTGT CTCGAAGCCT 2400
GAGGCACTGC TATTTCCAGT AAGCGACTGG CATCCCGAG ACCAGCCTCA GAAGGAGGAG 2460
GACCTGGGAC CTGCTGTCCA CCAATGTCTAT GAGCTCATCA ACCAAGSCCC CAGCTCCATT 2520
AGCCAGGGTG ACTGTGAACT CAGCTGTCCC CAGGCTCTGG AAGGTGAGCA GCTCCTATAT 2580
GTGACAGAGG TTAGCGGACT CACTGCAACC ACCAATCACC CCATTAAACC AAAGGGCCTG 2640
GAGTTGGATC CCGAGGGTTC CCTGCAACAC CAGCAAAAAC GGGAAAGCTCC AAGCCGAGC 2700
TCTGCTTCCI CGGGACCTCA GATCCTGAAA TGCCCGAGG CTGAGTGTGT CAGGCTGCGC 2760
TGTGAGCTCG GGCCTCTGCA CCAACAAGAG AGCCAAAGTC TGCAATTGCA TTTCCGAGTC 2820
TGGGCCAAGA TCTTTCTGCA GCGGGAGCAC CAGCCATTTA GCTTGCAGTG TGAGGCTGTG 2880
TACAAAGCCC TGAAGATGCC CTACCGAATC CTGCTCGGCG AGCTGCCCCA AAAAGAGCGT 2940
CAGGTGGCCA CAGCTGTGCA ATGGAACCAAG GCAGAAGGCA GCTATGGGCT CCCACTGTGG 3000
ATCATCATCC TAGCCATCCT GTTTGGCCTC CTGCTCCTAG GTCTACTCAT CTACATCCTC 3060
TACAAGCTTG GATTCTTCAA ACGCTCCCTC CCATATGGCA CCGCATGGA AAAAGCTCAG 3120
CTCAAGCCTC CAGCAACCTC TGATGCTTGA

```

Seq ID NO: 4 Protein sequence
Protein Accession #: NP_002196.1

50
55
60
65
70

```

1 11 21 31 41 51
| | | | |
MGRTPESPL HAVQLEWGR RRPFLPLLL LLLPPPRVG GNLDAEAPA VLSGPPGSPF 60
GFSEVYRPG TDGVSVLVLA PKANTSQRGV LQGGAVLCP WGRSPQCTP IEPDSXGRL 120
LESSLSBSEG BEPVEYKSLQ WFGATVRAHG SSILACAPLY SWRTEKEPLS DPVGTCTYST 180
DNFRIILEYA PCRSDFSWAA GQGYCQGGFS AEPTKTRGVV LGGPGSYFHO GQILSATYEQ 240
IARSTYPEYL INLVQQLQT HQASSIYDSS YLGSYVAVGE FSGDDTDFV AGVPKGNLTY 300
GYVTILNGSD IRSLYNFBGS QMASYFGYAV AATDVNGDL DLLLVGAPLL MDETDFGRFQ 360
EYGRVYVYLQ HPAGIEPTPT LTLTGHDFEG RFGSLLTFLG DLDQDGYNDV AIGAPFQGET 420
QQGVVVFVPG GPGGLSGKPS QVLQFLMAAS HTPDFFGSAL RGRDLIDNG YPDLTVGSFG 480
VDKAVVYRGR PIVSASASLT IFPAMFNPER RSCSLGNPV ACINLSFCLN ASGKHVADSI 540
GFTVELQLDW QKQGGVRRRA LFLASRQATL TQTLIIQNGA REDCREMKIY LRNSSEPRDX 600
LSPHIALMF SLDQPAFVDS EGLRPALHYO SKSRIEDKAO ILDDCGEDNI CYEDLQLEVF 660
GBOHMYVLGD KNALNLTFFA QNVGEGGAYE AELRVTAPE AEYSGIVRHP GNFSSLSQDY 720
FAVNSQRLIV CDLGNPMKAG ASLWGLRFT VPHLEDTKKT IQDFQILSK NLNNSQSDVV 780
SFRLSVEAQA QVTINGVSKP EAVLFPVSDW HPRDQPKER DLGPAVHHVY ELINQGPSII 840
SQGVLELSCP QALEGQQLLY VTRVTGLNCT TNEPINPKGL ELDPEGLLHE QQKREAPSR 900
SASSGPQILK CPEABCFRLR CELGPLHQOE QSLSQLHFRV WAKTFLOREH QPFSLQCEAV 960
YKALKMPYRI LPRQLPKOKR QVATAVQNTK AEGSYGVPLN IITLAILFLG LLLGLLIYTL 1020
YKLGFFKRSL PYGTAMEKAQ LKPPATSDA

```

Seq ID NO: 5 DNA sequence
Nucleic Acid Accession #: NM_002211.1
Coding sequence: 1..2397

75
80

```

1 11 21 31 41 51
| | | | |
ATGAATTAC AACCAATTT CTGGATTGGA CTGATCAGTT CAGTTTGTCT TGTGTTGCT 60
CAACAGATG AAAATAGATG TTAAAGACA AATGCCAAAT CATGTGGAGA ATGTATACAA 120
GCAGGGCCAA ATGTGTGGTG GTGCACAAAT TCAACATTTT TACAGGAAGG AATGCCCTACT 180
TCTGCACGAT GTGATGATTT AGAAGCCCTA AAAAAGAGGG GTTGCCCTCC AGATGACATA 240
GAAATATCCA GAGGCTCCAA AGATATAAAG AAAAATAAAA ATGTAAACCA CCGTAGCAAA 300
GGAACAGCAG AGAAGCTCAA GCCAGAGGAT ATTACTCAGA TCCACACACA GCAGTTGGTT 360
TTGCGATTAG GATACGGGGA GCCACAGACA TTACATTAA AATTCAAGAG AGCTGAAGAC 420
TATCCCATTT ACCCTACTA CTTATGGAC CTGCTTACT CAATGAAAGA CGATTGGAG 480
AATGTAAGAA GTCTTGAAC AGATCTGATG AATGAAATGA GGAGGATTAC TTGGGACTTC 540

```

5
10
15
20
25
30

```

AGAATGGGAT TTGGCTCATT TGTGGAAAAG ACTGTGATGC CTTACATTAG CACAACACCA 600
GCTAAGCTCA GGAACCCCTG CACAAGTGAA CAGAACTGCA CCAGCCCATT TAGCTACAAA 660
AATGTGCTCA GTCTTACTAA TAAAGGAGAA GTATTTAATG AACTTGTGGG AAAACAGCGC 720
ATATCTGGAA AITTGGATTG TCCAGAAAGT GGTTCGATG CCATCATGCA AGTTGCAGTT 780
TGTGATCAC TGAATGGCTG GAGGAATGTT ACACGGCTGC TGGTGTITTC CACAGATGCC 840
GGGTTCCTCT TTGCTGGAGA TGGGAAACTT GGTGGCATTG TTTTACCATA TGATGGACAA 900
TGTCACTTGG AAAATAATAT GTACACAATG AGCCATTATT ATGATTATCC TTCTATTGCT 960
CACCTTGTCC AGAAACTGAG TGAAATAAAT ATTCAGACAA TTTTTCAGT TACTGAAGAA 1020
TTTCAGCTCG TTTACAGGA GCTGAAAAAC TTGATCCCTA AGTCAGCAGT AGGAACATTA 1080
TCTGCAAAAT CTAGCAATGT AATTCACTTG ATCATTGATG CATACAATTC CCTTTCCTCA 1140
GAAGTCATTT TGGAAAACGG CAAATTGTCA GAAGGAGTAA CAATAAGTTA CAAATCTTAC 1200
TGCAGAACGG GGTGTAATGG AACAGGGGAA AATGGAAGAA AATGTTCCAA TATTTCCATT 1260
GGAGATGAGG TTCAATTGGA AATTAGCATA ACTTCAAATA AGTGTCCAAA AAAGGATTCT 1320
GACAGCTTTA AAATTAGGCC TCTGGGCTTT ACAGGAGGAG TAGAGGTAT TCTTCAGTAC 1380
ATCTGTGAAT GTGATGCCA AAGCGAAGGC ATCCCTGAAA GTCCCAAGTG TCATGAAGGA 1440
AATGGGACAT TTGAGTGTGG CGCGTGCAGG TGCAATGAAG GCGGTGTGGG TAGACATTGT 1500
GAATGCAGCA CAGATGAAGT TAACAGTGAA GACATGGATG CTTACTGCAG GAAGGAAAAC 1560
AGTTCAGAAA TCTGCAATTA CAATGGAGAG TCCGTCTGCG GACAGTGTGT TTGTAGGAAG 1620
AGGGATAATA CAATGAATAT TTATTCTGGC AAATCTCTGG AGTGTGATAA TTCAACTGT 1680
GATAGATCCA ATGCTTAAT TTGTGGAGGA AATGTTGTTT GCAAGTGTGG TGTGTGAG 1740
TGCAACCCCA ACTACACTGG CAGTGCATGT GACTGTTCTT TGGTACTAG TACTTGTGAA 1800
GCCAGCAACG GACAGATCTG CAATGGCCGG GGCATCTGCG AGTGTGGTGT CTGTAAAGTT 1860
ACAGATCGGA AGTTTCAAGG GCAAACTGTG GAGATGTGTC AGACCTGCTT TGGTGTCTGT 1920
GCTGAGCATA AAGATATGTT TCAGTGACGA GCTTCAATA AAGGAGAAAA GAAAGACACA 1980
TGCAACAGG AATGTTCTTA TTTTAACATT ACCAAGGTAG AAAGTCCGGA CAAATTACCC 2040
CAGCGGGTCC AACCTGATCC TGTGTCCCAT TGTAAAGGAG AGGATGTTGA CGACTGTTGG 2100
TTCTATTTTA CTAATCTAGT GAATGGGAAC AACGAGGTCA TGGTTCATGT TGTGGAGAA 2160
CCAGATGTGC CCACCTGCTC AGACATCAT CCAATTGTAG CTGGTGTGGT TGTCTGGAAT 2220
GTTCTTATGT CCCTTGCATT ACTGTGATA TGAAGCTTT TAATGATAAT TCATGACAGA 2280
AGGGAGTTTG CAAATTTTGA AAAGGAGAAA ATGGAATGCA AATGGGACAC GGGTGAAAA 2340
CCTATTTTATA AGAGTGCCGT AACCACTGTG GTCAATCCGA AGTATGAGGG AAAATGA

```

Seq ID NO: 6 Protein sequence
Protein Accession #: NP_002202.1

35
40
45
50

```

1 11 21 31 41 51
MNLQIPFWIG LISSVCCVFA QTDENRCLKA NAKSCGECIQ AGPNCGWCTN STFLQBMPT 60
SARCDLEAL KKKGCPPIID ENPRGSKDIK KKNVITNRK GTAELKLPED ITQIQPQVLV 120
LRLRSGEPQT PTLKFKRAED YPIDLYLMD LSYSMKDDLE NVKSLGTDLM NEMRRITSDP 180
RIGFGSFVEK TVMPYISTTP AKLRNPCTSE QNCTSPFPGYK NVLSL/INKGE VFNELVCKQR 240
ISGNLDSPEG GFDALMQYAV CGSLIGHENV TRLLVFTDA GFHPAGDGKL GGIVLEMDGQ 300
CHLENNMYTM SHYVDYPSIA HLVDKLSENN IQTIFAVTEE FQPVYKELKN LI PKSAVGTL 360
SAMSSNVIGL ILDAYNSLSS EVILENGLKS EGVITISYKSY CKNGVNGTGE NGRKCSNISI 420
GDEVQFEISI TSNKCPKKDS DSPKIRPLGF TEEVEVILQY ICECEQSEG IPSPKCHBG 480
NGTFEGACR QNEGRVGRHC ECSTDEVNSE DMDAYCRKEN SSEICSNNGE CVGQCVCVRK 540
RDNTBIYSG KFCECDNFNC DRNGLICGG NGVCKCRVCE CNFNYTGSAC DCSLDTSTCE 600
ASNGQICNCR GICECGVCKR TDPKFQGTQC EMCOTCLGVC AEHKECVQCR APNKGKKDT 660
CTQECSEYFI TKVSRSDKLF QFVQFPVSH CKERDVDDCW FYFTYSVNGN NEVMVHEVEN 720
PECTTGEDI PIVAGVAGI VLIGLALLLI WKLLMIHIDR REFAPFEREK MNARNDTGEN 780
PIYKSAVTTV VNPKYEGK

```

Seq ID NO: 7 DNA sequence
Nucleic Acid Accession #: NM_002425
Coding sequence: 26..1453

55
60
65
70
75
80

```

1 11 21 31 41 51
AAAGAGGTA AGGGCAGTGA GAATGATGCA TCTTGCAATC CTTGTGCTGT TGTGTCTGCC 60
AGTCTGCTCT GCCTATCTCT TBAATGGGGC AGCAAAAGAG GAGGACTCCA ACRAGGATCT 120
TGCCACAGCA TACCTAGAAA AGTACTACAA CCTCGAAAAG GATGTGAAC AGTTTAGAAG 180
AAAGGACAGT AATCTCATTT TTAATAAAT CCAAGGAATG CAGAAGTCC TTGGGTGGA 240
GGTGACAGGG AAGCTAGACA CTGACACTCT GGAGGTGATG CGCAAGCCCA GGTGTGAGT 300
TCTTGACGTT GGTCACTTCA GCTCCTTTCC TGGCATGCCG AAGTGGAGGA AAACCCACCT 360
TACATACAGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTGT ATTCTGCCAT 420
TGAGAAAGCT CTGAAAGTCT GGGAAAGAGT GACTCCACTC ACATCTCTCA GGCTGTATGA 480
AGGAGAGGCT GATATAATGA TCTCTTTGCG AGTAAAGAA CATGGAGACT TTTACTCTTT 540
TGAEGGCCCA GGACACAGIT TGGCTCATGC CTACCCACCT GGACCTGGEC TTTATGGAGA 600
TATTCACCTT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATTCTT 660
CGTGTCTGCT CATGACCTGG GCGACTCCCT GGGGCTCTTT CACTCAGCCA ACACGTGAAGC 720
TTTGATGTAC CCRCTCTACA ACTCATTCAC AGAGCTCGCC CAGTTCGCTT TTTGCAAGA 780
TGATGTGAAT GGCATTCAGT CTCTCTACGG ACCTCCCTCT GCCTCTACTG AGGAACCTCT 840
GGTGCCACAG AATCTGGTTC CTTCGGGATC TGAGATGCCA GCCAAGTGTG ATCCTGCTTT 900
GTCTCTCGAT GCCATCAGCA CTCTGAGGGG AGAATATCTG TTCTTTAAG ACAGATATT 960
TTGGCGAAGA TCCACTGGA ACCCTGAACC TGAATTCAT TTGATTCTG CATTTTGGCC 1020
CTCTCTTCCA TCTATTGAG ATGCTGCATA TGAAGTTAAC AGCAGGGACA CCGTTTTTAT 1080
TTTTAAAGGA AATGAGTTCT GGGCCATCAG AGGAATAGAG GTACAAGCAG GTTATCCAG 1140
AGGCATCTG ACCCTGGGTT TTCTTCCAAC CATAAGGAAA ATGTGTGAG CAGTCTCTGA 1200
CAAGGAAAAA AAGAAAACAT ACTTCTTTGC AGCGGACAAA TACTGGAGAT TTGATGAAAA 1260
TAGCCAGTCC ATGGACCAAG GCTTCCCTAG ACTAATAGCT GATGACTTTC CAGGAGTTGA 1320
GCCTAAGGTT GATGCTGTAT TACAGGCATT TGGATTTTTC TACTTCTTCA GTGGATCATC 1380
ACAGTTTGA TTTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAG GTAACAGCTG 1440
GTTACATTGC TAGGCGAGAT AGGGGGAAGA CAGATATGGG TGTTTTAAAT AAATCTAATA 1500
ATTATTCATC TAATGTATTA TGAGCCAAAA TGGTTAATTT TTCTGCTGCT TTCTGTGACT 1560

```


GAAGAAGATG AGCCTTGCGG ATATCTGCAT GTGTCATGAA GAATGTTTCT GGAATTCCTC 1620
 ACTTGCTTTT GATTTCACCT GAACAGAAAT AAGAAATACT CATGTGCAAT AGGTGAGAGA 1680
 ATGTATTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTTATTTT GGGCCTGTC 1740
 CTT

Seq ID NO: 8 Protein sequence
 Protein Accession #: NP_002416

1	11	21	31	41	51	
MHLAFVLVLC	LPVCSAYPLS	GAAKERDSNK	DLAQQYLEKY	YNLEKDVVKQF	RRKDSNLIVK	60
KIQGMQKFLG	LEVITGRDLD	TLEVMRKPRC	GVPDVGHFSS	FPGMPKWRKT	HLTTRIVNYT	120
PDLPDRAVDS	AIEKALKVNE	EVPTLTF SRL	YEGEADIMIS	PAVKERHDFY	SFDGPGHSLA	180
HAYFFGPGGLY	GDIHFDDDEK	MTEDASGTNL	FLVAAHELGH	SLGLEHSAAT	EALMYPLVNS	240
PTELAQFRLS	QDDVNGIQSL	YGPFPASTE	PLVPTKSVPS	GSEMPAKCDP	ALSPDAISTL	300
RGEYLFFKDR	YFWRSHWNP	EPFHLISAF	WPSLPSYLD	AYEVNSRDTV	FIFKNEFWA	360
IRGNVQAGY	PRGHTHTGFP	PTIRKIDAAV	SDKSKKTYF	PAADKYWRFD	ENSQSMQGF	420
PRLLADDFG	VEPKVDVAVLQ	AFGFFYFSSG	SSQFEFDENA	RMVTHILKSN	SWLHC	

Seq ID NO: 9 DNA sequence
 Nucleic Acid Accession #: XM_058189.2
 Coding sequence: 169..774

1	11	21	31	41	51	
GAAGACCAGC	TCAGCTCTTC	AGTTGTGTAT	CATTGTCTAT	TGTTCTCCAA	ACAGTAAACC	60
AGTATTTCAC	ACTGAGATTG	TGCGCTGCGG	GTATATTCCA	ATTCCTCCGT	TCCTCATGAA	120
TATGAGTAGT	AGGCTCTGGA	CCCTGGAAAGT	GGTCTAAGC	AGGGCAAAAT	GGGCTCTCGG	180
AAGTGTGGAG	GCTGCTTAAG	TTGTTTGTCT	ATTCGCTCTG	CACCTTGGAG	TATAATCGTG	240
AAACATATTAT	TGTATTTCCT	GAATGGGCAA	ACTTCTATG	CATCCAGCAA	TAAACTCAAC	300
AACTACGTGT	GGTATTTTGA	AGGAATCTGT	TTCTCAGGCA	TCATGATGCT	TATAGTAACA	360
ACAGTCTCTC	TGGTACTGGA	GAATAATAAC	AACTATAAAT	GTTCGCGAG	TGAAAACCTG	420
AGCAAAAAT	ATGTGACACT	GCTGTCAATT	ATCTTTTCTT	CCCTCGGAAT	TGCTTTTCTT	480
GGATACGCGC	TGGTCACTCT	TGCTTGGGT	CTGTCTCAAG	GGCCATATTG	CCGCACCTCT	540
GATGGCTGGG	AGTATGCTTT	TGAAGGCAT	GCTGGACGTT	TCCTACAGA	TTCTAGCATA	600
TGGATTCAGT	GCCTGGAACC	TGCACATGTT	GTGGAGTGG	ACATCATTTT	ATTTTCCATT	660
CTCATTAACC	TCAGTGGGCT	TCAAGTGATC	ATCTGCTCA	TCAGAGTAGT	CATGCAACTA	720
TCCAAGATAC	TGTGTGGGAG	CTATTCAAGT	ATCTTCCAGC	CTGGAATCAT	TTGAATAAGG	780
ACAAAATGTT	TTCCATTATC	AAGACATGGC	CATCTATCTA	AATATTATAT	CAACTGTGTA	840
GACTTGAGGG	CAATATTGAA	ATGATGGTGC	TTTCTGCATT	TGGTGTATT	TTGTAATAAA	900
TTTGCGATCC	TCATGTCACA	TGCAAGTATA	CCACCTTCC	ATTAGATATG	TTTTTTAAGT	960
AATATGATC	AGAACTTCA	GAAATAGTTC	TGCCCTTTGA	TCAACAAAT	CCATTTCCTA	1020
GAATCTGTAC	TAGGGAAGTA	AATAAGATA	TGAGAGAAAC	CTTATGCAA	ATATGTATAT	1080
TGCAACATTA	TTAATATTTC	TGCAAAATG	GAAACACCCC	AAAATCTTAA	ACTCAGAGGA	1140
AGGATTAAAT	AAAGAGTGGT	ACATACGTGA	AATGTTTTCT	GATATTAATA	AAAAAATTAA	1200
ATAAAAAATA	AAGAGTACTA	CATGTTGTGA	AAA			

Seq ID NO: 10 Protein sequence
 Protein Accession #: XP_058189.1

1	11	21	31	41	51	
MGERKGGGCL	SCLLIFLALM	SIIVNILLVY	ENGQTSYASS	NKLINIVVWYF	EGICFSGIMM	60
LIVTTVLVLV	ENNNRYKCCQ	SENCSSKYVT	LLSILFSSLG	IAPSGYCLVI	SALGLVQGPY	120
CRITLDGNEVA	FEPTAGRLIT	DSSIIWICLE	PAHVVEWNII	LFSILITLBS	LQVILCLRV	180
VMQLSKILCG	SYSVIFQPGI	I				

Seq ID NO: 11 DNA sequence
 Nucleic Acid Accession #: NM_002421.2
 Coding sequence: 1..1409

1	11	21	31	41	51	
ATGCACAGCT	TTCTCCACT	GCTGCTGCTG	CTGTTCTGGG	GTGTGCTGTC	ACACAGCTTC	60
CCAGCGACTC	TAGAAACACA	AGAGCAAGAT	GTGGACTTAG	TCCAGAAATA	CCTGGAAAAA	120
TACTACAAAC	TGAAGAATGA	TGGGAGGCAA	GTGAAAAGC	GGAGAATAG	TGGCCAGATG	180
GTGAAAAAT	TGAAGCAAT	GCAGGAATTC	TTTGGGCTGA	AAGTGACTGG	GAAACCAAGT	240
GCTGAAACCC	TGAAGTGAT	GAAGCAGCCC	AGATGTGGAG	TGCTGATGT	GGCTCAGTTT	300
GTCCTCACTG	AGGGGAACCC	TGCTGGGAG	CAACACATC	TGACCTACAG	GATTGAAAAT	360
TACAGCCAG	ATTTGCCAG	AGCAGATGTG	GACCATGCCA	TTGAGAAAGC	CTTCCACTC	420
TGGAGTAATG	TCACACCTCT	GACATTCAAC	AAGTCTCTG	AGGCTCAGC	AGACATCATG	480
ATATCTTTTG	TCAGGGGAGA	TCATCGGAC	AACTCTCCTT	TTGATGGACC	TGGAGGAAAT	540
CTTGCTCATG	CTTTTCAACC	AGGCCAGGT	ATTGGAGGGG	ATGCTCATTT	TGATGAAGAT	600
GAAAGGTGGA	CCACAAATTT	CAGAGAGTAC	AACTTACATC	GTGTTCCGGC	TCATGAATCT	660
GGCCATTCTC	TTGACCTCTC	CCATTCTACT	GATATCGGGG	CTTTGATGTA	CCCTAGCTAC	720
ACCTTCAGTG	GTGATGTTCA	GCEAGCTCAG	GATGACATG	ATGGCATCCA	AGCCATATAT	780
GGACGTTCCC	AAAATCCTGT	CCAGCCATC	GGCCACAAA	CCCCAAAAGC	ATGTGACAGT	840
AAGCTAACCT	TTGATGCTAT	AACTACGATT	CGGGGAGAAG	TGATGTTCTT	TAAAGACAGA	900
TTTACATGCT	GCACAAATCC	CTTCTACCCG	GAAAGTTGAGC	TCAATTTCAT	TTCTGTTTTC	960
TGGCCACAC	TGCCAAATGG	GCTTGAAGCT	GCTTACGAAT	TTGCCGACAG	AGATGAAGTC	1020
CGGTTTTC	AAGGGAAATA	GTACTGGGCT	GTTCAGGGAC	AGAAATGTCT	ACACGGATAC	1080
CCCAAGGACA	TCTACAGCTC	CTTTGGCTTC	CCTAGAAGTC	TGAAGCATAT	CSATGCTGCT	1140
CTTTCTGAGG	AAACACTGG	AAAAACCTAC	TTCTTTGTTG	CTAACAAATA	CTGGAGGTAT	1200
GATGAATATA	AACGATCTAT	GGATCCAGGT	TATCCCAAAA	TGATAGCACA	TGACTTTCCT	1260

GGAATTGGCC ACAAGTTGA TGCAGTTTC ATGAAAGATG GATTTTTCIA TTTCTTTCAT 1320
GGAACAAGAC AATACAAAT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG

5 Seq ID NO: 12 Protein sequence
Protein Accession #: NP_002412.1

1 11 21 31 41 51
10 MHSFPPLLLL LFWGVVSESP PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLRQMGEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRNE QTHLYRIEN 120
YTPDLPRADV DHAIEKAPQL WSNVTPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180
LAHAFQPGFG IGGDAHFED EKWNNFREY NLHRVAHAEL GHSLGLSHST DIGALMYPST 240
15 TFSGDVQLAQ DDIDGIAIY GRSQNPVQPI GPQTFKACDS KLTFDALTTI RGEVMFFKDR 300
PYMRTNPFYP EVELNFISSF WPQLPNGLA AYEFADRDEV RPFKGNKYWA VQGQNVLRGY 360
PKDIYSSFGF PRYVKHIDAA LSEENTGKTY FFWANKYWRV DEYKRSMDPG YPKMIADHFP 420
GIGHKVDVAF MKDGFFFFFH GTRQYKFDPK TKRILTLQKA NSWFNCRKN

20 Seq ID NO: 13 DNA sequence
Nucleic Acid Accession #: NM_002421.2
Coding sequence: 1..1409

1 11 21 31 41 51
25 ATGCACAGCT TTCCTCCACT GCTGCTGCTG CTGTTCTGGG GTGTGSGTGC ACACAGCTTC 60
CCAGCCAGCTC TAGAAACACA AGAGCAAGAT GTGGACTTAG TCCAGAAATA CCTGGAAAAA 120
TACTACAACC TGAAGAATGA TGGGAGGCAA GTTGAAGAGC GGAGAAATAG TGGCCCACTG 180
30 GTTGAAGAAAT TGAAGCAATC GCAGGAATTC TTTGGGCTGA AASTGACTGG GAAACCAAGT 240
GCTGAAGACC TGAAGGTGAT GAAGCAGCCC AGATGTGGAG TGCCGTGATGT GGCCTCAGTT 300
GTCTCTACTG AGGGGAACCC TCGCTGGGAG CAACACATC TGACCTACAG GATTGAAAAAT 360
TACAGCCAGC ATTTGCCAAG AGCAGATGTG GACCATGCCA TTGAGAAAGC CTTCACATC 420
TGGAGTAATG TCACACCTCT GACATTCAAC AAGGTCTCTG AGGCTCAAGC AGACATCATG 480
ATATCTTTTG TCAGGGGAGA TCATCGGGAC AACTCTCCTT TTGATGGACC TGGAGGAAAT 540
CTTGCTCATG CTTTCAACCC AGGCCAGGT ATTGGAGGGG ATGCTCATTT TGATGAAGAT 600
35 GAAAGGTGGA CCAACAATT CAGAGAGTAC AACTATATC GTGTGSGGC TCATGCCCTC 660
GGCCATCTCT TTGGACTCTC CCATTCTACT GATATCGGGG CTTTGATGTA CCTAGCTAC 720
ACCTTCAGTG GTATGTTTCA GCTAGCTCAG GATGACATG ATGGCATCCA AGCCATATAT 780
GGAAGTTCCC AAAATCCGAT CCAGCCCATC GGGCCACAAA CCCCAGAGC ATGTGACAGT 840
AAGCTAACCT TTGATGCTAT AACTACGATY CGGGGAGAGG TGATGTTCTT TAAAGACAGA 900
40 TTCTACATGC GCACAAATCC CTTCTACCCG GAAGTTGAGC TCAATTTCAT TCTGTTTTTC 960
TGCCCAACAC AGGGAATGG GCTTGAAGCT GCTACGAAT TTGCCGACAG AGATGAAGTC 1020
CGGTTTTTCA AAGGGAATAA GTACTGGGCT GTTCAGGGAC AGAATGTGCT ACACGGATAC 1080
CCCAAGGACA TCTACAGCTC CTTTGGCTTC CCTAGAACTG TGAAGCATAT CGATGCTGCT 1140
45 CTTTCTGAGG AAAACACTGG AAAAACTTAC TTCTTTGTTG CTAACAAATA CTGAGGATAT 1200
GATGAATATA AACGATCTAT GGATCCAGGT TATCCCAAAA TGATAGCACA TGACTTTCTT 1260
GGAATTGGCC ACAAAGTTGA TGCAGTTTTC ATGAAAGATG GATTTTCTTA TTTCTTTCAT 1320
GGAACAAGAC AATACAAAT TGATCCTAAA ACGAAGAGAA TTTTGACTCT CCAGAAAGCT 1380
AATAGCTGGT TCAACTGCAG GAAAAATTAG

50 Seq ID NO: 14 Protein sequence
Protein Accession #: NP_002412.1

1 11 21 31 41 51
55 MHSFPPLLLL LFWGVVSESP PATLETQEQD VDLVQKYLEK YYNLKNDGRQ VEKRRNSGPV 60
VEKLRQMGEF FGLKVTGKPD AETLKVMKQP RCGVPDVAQF VLTEGNPRNE QTHLYRIEN 120
YTPDLPRADV DHAIEKAPQL WSNVTPLTFT KVSEGGADIM ISFVRGDHRD NSPFDGPGGN 180
60 LAHAFQPGFG IGGDAHFED EKWNNFREY NLHRVAHAEL GHSLGLSHST DIGALMYPST 240
TFSGDVQLAQ DDIDGIAIY GRSQNPVQPI GPQTFKACDS KLTFDALTTI RGEVMFFKDR 300
PYMRTNPFYP EVELNFISSF WPQLPNGLA AYEFADRDEV RPFKGNKYWA VQGQNVLRGY 360
PKDIYSSFGF PRYVKHIDAA LSEENTGKTY FFWANKYWRV DEYKRSMDPG YPKMIADHFP 420
GIGHKVDVAF MKDGFFFFFH GTRQYKFDEK TKRILTLQKA NSWFNCRKN

65 Seq ID NO: 15 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 141..1580

1 11 21 31 41 51
70 TCTGCGTGTG CCGGGGCTAG GGGCTGGAAG TCCCTGGCTCT AGTTGCACCT CGGAAGGAAA 60
AGGCAACACAG AGGAGGGGAG GCGTCTTAGG ACTGCGTGA TCCAGAGCAC TTTCTCGGC 120
75 CTCACAGGC CTGTGTGCTC ATGGGTTCCG CCGCCGCCCC GGAGGGAGCG CTGGGCTAGG 180
TCCGCGAGTT CACTGCGCAC TCCTCCGACG TCGTGGGCAA CCTCAAGAG CTGCGCTGCT 240
GCGGGATCCT CACTGACGTC AGCGTCTGAG TTGGCGGGCA ACCCTTCAGA GCACACAGG 300
CAGTTCTCAT GCGCTGCACT GCGTCTCTCT ATTCAAATTT CCGGGGCGGT GCGGGAGTGG 360
GGGTGGAGCT GCTCTCTCTG CCGGGGCGTC CCGAAGGAG AGGCTTCGCT CCTCTATTGG 420
80 ACTTCATGTA CACTTCGCGC CTGCGGCTCT CTCAGGCCAC TGCACTAGCA GTCTAGCGG 480
CGCGCACTA TTGTCAGATG GAGCACGTGG TCCAGGCATG CCACGCTTC ATCCAGGCCA 540
GCTATGAACC TCTGGGCATC TCCCTGCGCC CCTTGGAGAG AGAACCCCCA ACACCCCCAA 600
CGGCCCTTCC ACCAGGTAGT CCGAGGCGCT CCGAAGGACA CCGAGCCCCA CCTACTGAAT 660
CTGGAAGCTG CAGTCAAGGC CCGCCAGTTC CAGCCAGCCC TGAACCCAG GCCTGCAACT 720
GGAAAAAGTA CAGTACATC GTGCTAAACT CTCAGGCTTC CCAAGCAGGG AGCTTGGTCTG 780

5	GGGAGAGAAG TTCTGGTCAA CCTTGCCTCC AAGCCAGGCT CCCCAGTGGG GACGAGGCCT 840
	CCAGCAGCAG CAGCAGCAGC AGCAGCAGCA GTGAAGAAGG ACCCATTCCT GGTCCCCAGA 900
	GCAGGCTCTC TCCAACTGCT GCACTGTGC AGTTCAAATG TGGGGCTCCA GCCAGTACCC 960
	CCTACCTCCT CACATCCAGG GCTCAAGACA CCTCTGATC ACCCTCTGAA CGGGCTCGTC 1020
	CACIACCAGG AAGTGAATTT TTCAGCTGCC AGAAGCTGTA GGCTGTGGCA GGGTGTCTAT 1080
	CGGGGCTGGA CTCTTGGTTT CCTGGGGACG AAGACAAACC CTATAAGTGT CAGCTGTGCC 1140
	GGTCTTCGTT CCGCTACAAAG GGCAACCTTG CCAGTCATCG TACAGTGCAC ACAGGGGAAA 1200
	AGCCTTACCA CTGCTCAATC TGGGGAGCCC GTTTTAAACG SCCAGCAAAC CTGAAAACGC 1260
10	ACAGCCGCAT CCATTCGGGA GAGAAGCCGT ATAAGTGTGA GACGTGCGGC TCGCGCTTTG 1320
	TACAGTGGC ACATCTGCGG GCGCACGTGC TGATCCACAC CGGGGAGAG CCCTACCCTT 1380
	GCCTTACCTG CGGAACCCGC TTCGCCCACC TGCAGACCTT CAAGAGCCAC GTTCGCATCC 1440
	ACACCGGAGA GAAGCCTTAC CACTGCGACC CCTGTGGCCT GCATTTCGG CACAAGAGTC 1500
	AACCTGCGCT GCATCTGCGC CAGAAACACG GAGCTGCTAC CAACACCAA GTGCACTACC 1560
15	ACATTCTCGG GGGGCCCTAG CTGAGCGCAG GCCCAGGCC CACTGTCTTC CTGGGGTGG 1620
	GAAGCTGCA GGGCCAGGCC TTGCTTCCCT ATCAGGCTTG GGCATAGGGG TGTGCCAGGC 1680
	CACCTTGGTA GTGATAAATG CCACCCCTCT AATTTCCTAC TGGGGAGAGC AGGGGTGGCA 1740
	GATCTCGGCT AGATCTGCCT CTGTTTGTCT GGTCAAAACC TCTTCCCAC AAGCCAGATT 1800
	GTTCCTGAGG CAGAGCTAG CTAGGGGCTG GGAAGGGGA GAGATTGGAG TCCTGGTCTC 1860
20	CCTAAGGAA TAGCCCTCCA CCTGTGGCCC CCATTGCAT CAGTTATCT GTAAATATAA 1920
	TTTATTGAGG CCTTGGGGTG GCACCGGGGC CTTCAATCGA TGCATTTC CACTCCCTCC 1980
	TTCCACAAGT TGTATAGTGT TAGACCTTCT AGAAATGGAT ACAGACATT CTCTGTCTCT 2040
	CAGAGATTAC TAGCCCTTGG CTCTCTCGTT TGGCTTGGGT ATTTTATATT ATTCTGTCA 2100
	TAACTTTTAT CTTTAGAATT GTTCTTTCTC CTGTTTGTCT CTTGTTAGT TTGTTTAAAA 2160
25	TGGAAAAGG GGTCTCTGCT GTTCTGCCCC TGTAAATCTA GGTCTGGAAC CTTTATTGT 2220
	TCTAGGGCAG CTCTGGGAAC ATGCGGGATT GTGGAATTGG GTCAAGAAC CTCTCTGGTA 2280
	TTCTGGATGT TGTAGGTTCT CTAGCAGTCT AGAAATGGAT ACAGACATT CTCTGTCTCT 2340
	CAAGGGTGAT AGGAACCAT ATGTTGAGCC CAAAATGGAA GTAAATAATA ATGCCTCCTG 2400
	GAGGCTGTGG GTGTGGGGGA TTCTGTATCT GGATTCCTGA TCACTCCAAC TGGAGGCTGT 2460
30	GGGTGTGGGG GATTCTGTAT CTGGATTCCG TATCACTCCA AGTGGAGGCT GGCAGGTTTT 2520
	TCTGCAAGAT GGTCCAGAA CTAAATGTCT CCATTAACTI GGTCACTTGG GTTGTGCTCT 2580
	GCTGTATCCA TCTATAGTGG TAGAGACCCA CCAGGCTCA AGTGGAGTCC ATCATCTCTC 2640
	CACGGGGGCC TGTCTTAGC ACTGAGTTGA TCGCTCCATG GGGGAGAGAT CAGACATTCC 2700
	TTATCAGAGA TTATCTGACT TTCTCTGACT CTGCCAGTC TCTATGAATG TTATGGCTTA 2760
35	GGGAAGAAATC ATGAAACTCT TTAGCTTGAT TAGATGGTAA ACAGTGTAA CCCATCTCTT 2820
	ACTACAGAG CATATGGGTT TGAATGTAC CTGGGGTTCT CTCTATTGAG TTGAGCCCTT 2880
	TCCTCTTTA GTGGGTTTGT GACATCTCTT GCAAGTGTCT CAGATGCCAG AACCTTCTTT 2940
	TCCTCTAGAA GGGATGGTGC TTGGTAACTT TACCTTTTAA AAGCTGGGTC TGTGACCTGG 3000
	TCCTCCATC CCTGCATTCC TGTCTGGAAC CAGTGAATGC ATTAGAACCT TCCATAGGAA 3060
40	AAGAAAAGGG GCTGAGTTCC ATTCTGGGTT TGTCTGAGTT TGGTGGGAT TATTGTTGGC 3120
	ATTACAGATG TAAAGATTG ACTAGCCCAT AGGCCAAGG CCTGTCTAG TTGACCAAGT 3180
	TTCAAGTAGG ATTAAGAGGT TGGTTGAGGG GTGCAGTTTC TGGTGTAGGC CAGGTAGGTA 3240
	GAAGATGAGG AACAGGGTTG CCTCTTGCT GGGTGGAGTC TCTGAATGT TAGAAGAGC 3300
	GCTGAAGCCT TGATTGATAG TTCTGCCCTT TGTGCCCCTG GGGCTTATCT GATTATGGGA 3360
45	CGAGGGTAGA AGTAAGAAG CACTTTTGA TTTTGGGGT AGAATTCAA CAATAAGTCA 3420
	GTCTAGTGG CTGTCGCCCT GGGACTAGTG AGAAGCTAC TCTTCTCCCT CTTCCTCTCT 3480
	TCTCCCCATG CCCCCACTGC AGAATTAAG AAGGAAGAG GGAAGGCGGA GGAGTCTATA 3540
	AGAAGGAATC ATGATTCTTA TTATAGCAGT TGGATGGGCA GGTGGAGAAT GCCTGGGGGT 3600
50	AGAAATGTTA GATCTTGCAG CATCAGATCC TTGGAATAAA GAAGCCTCTC TGYGCWRAAA 3660
	AAAAAAAAA AAAAAA

Seq ID NO: 16 Protein sequence
Protein Accession #: FGENESH predicted

55	1	11	21	31	41	51	
	MGSPAAPEGA	LGYYREFTRH	SSDVLGNLNE	LRLRGILTDV	TLLVGGQPLR	AHKAVLIACS	60
	GFYFVIFRGR	AGVGVVDVLS	PGGPEARFPA	FLDDPMYTSR	LRLSPATAFA	VLAATYLMQ	120
	EHVVQACHRF	IQASYEPLGI	SLRPLEAEPF	TPPTAPPFGS	PERSEGHDPD	PTESRCSQG	180
60	PFSPASPDFX	ACNWKYKYI	VLNQASQAF	SLVGERSSGQ	PCFQARLPFG	DEASSSSSSS	240
	SSSSSESPDP	GPQSRLSPTA	ATVQFKCGAF	ASTPYLLTSQ	AQDTSQSPSE	RARPLPSEPF	300
	FSCQNCBAVA	GCSGLDLSLV	PGEDKPYKFC	QLCRSSPFYK	ENLASHRTVH	TGEKPYHCST	360
	CGARFNRPAN	LKTHSRHS9G	EKPYKCTCGF	SRPVQVAHLR	AEVLITGEK	PYPCTCGTR	420
	FRLQLTKSE	VRHETGSKPY	HCDPCGLHFR	HKSQLRLHLR	QKHGAATNFK	VHYEILGGP	

Seq ID NO: 17 DNA sequence
Nucleic Acid Accession #: XM_039209
Coding sequence: 1..2049

70	1	11	21	31	41	51	
	ATGCTGAAGA	TGCTCTCCTT	TAAGCTGCTG	CTGCTGGCCG	TGGCTCTGGG	CTTCTTTGAA	60
	GGAGATGCTA	AGTTTGGGGA	AAGAAACGAA	GGGAGCGGAG	CAGGAGGAG	AAGGTGCCCT	120
	AATGGGAACC	CCCCGAGAGC	CCTGAAAAGG	AGBACAGGA	GGATGATGTC	CCAGCTGGAG	180
75	CTGCTGAATG	GGGGAGAGAT	GCTGTGCGGT	GGCCTTACCC	CTGGCTGTGC	CTGCTGCCCT	240
	CGGAGTGACA	GCCCGGGGCT	AGGGCGCCTG	GAGAAATAAG	TATTTTCTGT	TACCAACAAC	300
	ACAGAAATGT	GGAAGTTACT	GGAGGAAATC	AAATGTGCAC	TTTGCTCTCC	ACATTCTCAA	360
	AGCCTGTTC	ACTCACCTGA	GAGAGAAGTC	TTGGAAGAG	ACCTAGTACT	TCCTCTGCTC	420
	TGCAAGAACT	ATTGCAAGA	ATTCTTTTAC	ACTTGCCGAG	GCCATATTCC	AGGTTTCCTT	480
	CAACAACTG	CGGATGAGTT	TGCTTTTAC	TATGCAAGAA	AAGATGTGGG	GTTGTGCTTT	540
80	CCAGATTTTC	CAAGAAAACA	AGTCAGAGGA	CCAGCATCTA	ACTACTTGGG	CCAGATGGAA	600
	GAATATGACA	AAGTGGAGA	GATCAGCAGA	AAGCACAAC	ACAACTGCTT	CTGTATTGAG	660
	GAGGTTGTGA	GTGGGCTGCG	CGACCCCGTT	GGTGCCCTGC	ATAGTGGGGA	TGGCTGCGAA	720
	CGTCTCTTCA	TTCTGGAATA	AGAAGTTTAT	GTGAAGATAC	TTACCCCTGA	AGGAGAAATT	780
	TTCAAGGAGC	CTTATTGGA	CATTACAAA	CTTGTTCAAA	GCGGAATAAA	GGGAGAGAGAT	840

5
10
15
20

```

GAAAGAGGAC TGCTAAGCCT CGCATTCAT CCCAATTACA AGAAAAATGG AAAGTTGTAT 900
GTGTCTCTATA CCACCAACCA AGAACGGTGG GCTATCGGGC CTCATGADCA CATTTCTTAGG 960
GTTGTGGAAAT ACACAGTATC CAGAAAAAAT CCACACCAAG TTGATTGAG AACAGCCAGA 1020
GTCTTTCTTG AAGTTGCAGA ACTCCACAGA AAGCATCTGG GAGGACAACT GCTCTTTGGC 1080
CCTGACGGCT TTTTGTACAT CATTCCTGGT GATGGGATGA TTACACTGGA TGATATGGAA 1140
GAAATGGATG GGTAAAGTGA TTTACAGGC TCAGTGCTAC GGCTGGATGT GGACACAGAC 1200
ATGTGTAACG TGCCTTATTC CATACCAAGG AGCAACCCAC ACTTCAACAG CACCAACCAG 1260
CCCCCGAAG TGTTTGCTCA TGGGCTCCAC GATCCAGGCA GATGTGCTGT GGATAGACAT 1320
CCCACTGATA TAAACATCAA TTTAACGATA CTGTGTTGAG ACTCCAATGG AAAAAACAGA 1380
TCATCAGCCA GAATCTACA GATAATAAAG GGGAAAGATT ATGAAAGTGA GCCATCACTT 1440
TTAGAATTCA AGCCATTGAG TAATGGTCTT TTGGTTGGTG GATTGTGATA CCGGGGCTGC 1500
CAGTCAGAAA GATTGTATGG AAGCTACGTG TTTGGAGATC GTAATGGGAA TTTCTTAAC 1560
CTCCAGCAAA GTCCTGTGAC AAAGCAGTGG CAAGAAAAAC CACTCTGTCT CGGCACTAGT 1620
GGGTCTGTA GAGGCTACTT TTCCGGTCAC ATCTGGGAT TTGGAGAGA TGAACTAGGT 1680
GAAGTTTACA TTTTATCAA CAGTAAAAGT ATGACCCAGA CTCACATGG AAAACTCTAC 1740
AAAATTGTAG ATCCCAAAAG ACCTTTAATG CCTGAGGAAT GCAGAGCCAC GGTACACCT 1800
GCACAGACAC TGACTTCAGA GTGCTCCAGG CTCTGTGAA ACGGCTACTG CACCCCAACG 1860
GGAAAGTGCT GCTGAGTCC AGGCTGGGAG GGGGACTTCT GCAGAACTGC AAAATGTGAG 1920
CCAGCATGTC GTCATGGAGG TGTCTGTGTT AGACCGAACA AGTGCTCTG TAAAAAGGA 1980
TATCTTGGTC CTCAATGTGA ACAAGTGGAC AGAAACATCC GCAGAGTGAC CAGGCGAGAC 2040
ATCACCTAG

```

25 Seq ID NO: 18 Protein sequence
Protein Accession #: XP_039209

30
35
40

```

1 11 21 31 41 51
| | | | |
MLKMLSFKLL LLAVALGFFE GDAKFERNE GSGARRRRL NGNPPKRLKR RDRRMSQLE 60
LLSGGEMLOG GFYPRISCCI RSDSPGLGRL ENKIFSVTNW TECGLLEET KCALCSPHSQ 120
SLFHSFEREV LKRDVLVPLL KDYCKEFFF TCRGHIPGFL QTTADEFCFY YARKDGLCF 180
PDFPRKQVRG PASNYLDQME EYDKVEEISR KHKENCFCIQ EVVSGLRQPV GALHSGDGSQ 240
RLFLLEKEGY VKILTPEGEI FKEPYLDIHK LVQSGIKGSD ERGLLSLAFH PMYKNGKLY 300
VBYTNGQER AIGEPDHILR VVEYTVSRKN PHQVDLRTAR VFLEVAEIER KHLGGQLLEF 360
EDGFLYIILG DGMITLDDME EMDGLEDPFG SVLRDLVDTD MCNVPYSIPR SFPHFNSTNQ 420
PPFVFAHGLH DPGRCVDRH PTDININLTI LCSDSNGKNR SSARILQIHK GKDYSESESL 480
LEKPFPSNGP LVGGFVYRGC QSERLYGSYV FGDNRGNFLT LQSPVTRQW QEKPLCLGTS 540
GSCRGYPSGH ILGFGDEDEL EYVILSSSKS MTQTHNGKLY KIVDPKRLPM PERCRATVQP 600
AQTLTSECSR LCRNGYCTPT GKCCSPGWE GDFCRTAKCE PACREGGVCV RPNKCLCKNG 660
YLGPOCEQVD RNIRRVTRAD IT

```

45 Seq ID NO: 19 DNA sequence
Nucleic Acid Accession #: NM_014331.2
Coding sequence: 1..1506

50
55
60
65
70
75
80

```

1 11 21 31 41 51
| | | | |
ATGGTCAGAA AGCCTGTTGT GTCCACCATC TCCAAAGGAG GTTACCTGCA GGGAAATGTT 60
AACGGGAGGC TGCTTCTCCT GGGCAACAAG GAGCCACCTG GGCAGAGAGAA AGTGACAGCTG 120
AAGAGGAAAG TCACTTTACT GAGGGGAGTC TCCATTATCA TTGGCACCAT CATTEGAGCA 180
GGAATCTTCA TCTCTCTCAA GGGCGTGTCT CAGAACACGG GCAGCGTGGG CATGTCTCTG 240
ACCATCTGGA CGGTGTGTGG GGTCTCTGTA CTATTGGAG CTTTGTCTTA TGCTGAATTG 300
GGAAACAATA TAAAGAAATC TGGAGGTGAT TACACATATA TTTTGGAGAT CTTTGGTGCA 360
TTACAGCTCT TTGTACAGAT CTGGGTGGAA CTCTCATATA TACGCGCTGC AGCTACTGCT 420
GTGATATCCC TGGCATTTGG ACCTACATTT CAGGAACCAT TTTTATTCCA AAGTGAAATC 480
CCTGAACCTG CGATCAAGCT CATTCAGCT GTGGGCATAA CTGTAGTGAT GGTCTCTAAT 540
AGCATGAGTG TCAGCTGGAG CGCCGGGATC CAGATTTTCT TAACCTTTTG CAAGCTCACA 600
GCAATTTGTA TAATATAGT CCTTGAAGT ATGCAGCTAA TTAAGGTGCA AACGCAGAAC 660
TTTAAAGACG CGTTTTCAGG AAGAGATTCA AGTATTACGC GGTGGCACT GGCTTTTAT 720
TATGGAGTGT ATGCATATGC TGGCTGGTTT TACCTCAACT TTGTACTGTA AGAAGTAGAA 780
AACCCTGAAA AAACCAATTC CCTTGCAATA TGTATATCCA TGGCCATTGT CACCATGGC 840
TATGTGCTGA CAATGTGGC CTACTTTAGC ACCATTATG CTGAGGAGCT GCTGCTTTCA 900
AATGCAGTGG CAGTGACCTT TTCTGAGCGG CTACTGGGAA ATTCTCTATT AGCAGTTCCG 960
ATCTTTGTTG CCTCTCCTG CTTTGGCTCC ATGAACGGTG GTGTGTTGCT TGCTCCAGG 1020
TTATCTATG TTGCGTCTCG AGAGGGTCAC CTCCAGAA TCCTCTCAT GATTGATGTC 1080
CGCAAGCACA CTCTCTTACC AGCTGTATT GTTTTGACC CTTTGACAA GATAATGCTC 1140
TTCTCTGAG ACCTCGACAG TCTTTTGAAT TTCTCAGT TTGCCAGGTG GCTTTTTATT 1200
GGGCTGGCAG TTGCTGGGCT GATTTATCTT CGATACAAAT GCCAGATAT GCATCGTCT 1260
TTCAAGGTGC CACGTTCAT CCCAGCTTGT TTTTCTTCA CATGCTCTT CATGGTTGCC 1320
CTTTCCTCT ATTCGGACC ATTAGTACA GGGATTGGCT TCGTCATCAC TCTGACTGGA 1380
GTCCCTGCGT ATTATCTCTT TATTATATGG GACAAGAAC CCAGGTGGTT TAGAATAATG 1440
TCAGAGAAA TAACCAGAAC ATTACAAATA ATACTGGAAG TTGTACCAGA AGAAGATAAG 1500
TATGAACATA ATGACATTGA GATCTTGGCA ATCTGCCCAA GGGAGACAC AAAATAGGGA 1560
TTTTTACTTC ATTTCTGAAA AGTCTAGAGA ATTACAACIT TGGTGATAAA CAAAGAGAGT 1620
CAGTTATTIT TATTATATA TTTAGCATA TTGCAACTAA TTTCTAAGAA ATTTAGTTAT 1680
AACTCTATGT AGTTATAGAA AGTGAATATG CAGTTATTCT ATGAGTCCGA CAATCTTGA 1740
GTCTCTGATA CCTACCTATT GGGGTAGGA GAAAGACTA GACAATTACT ATGTGGTCTAT 1800
TCTTACACAC ATATGTTAGC ACGCCAAAGA ACCTTCAAT TGAAGACTGA GATTTTCTG 1860
TATATATGGG TTTGTAAAG ATGTTTAC ACACACTAGA TGTCTATACT GTGAAAGATG 1920
TTTCAATTC TGAAAAAAG ATACATCAT GATTATGGCA AAGAGGAGAG AAAGAAATTT 1980
ATTTTACATF GACATTCAT TGCTTCCCT TAGATAACAA TTTAGATAAC AAACACTCAT 2040
GCTTTAATGG ATTATACCA GAGCACTTGT AACAAAGGTC AGTGGGATT GTTGAATACA 2100
TTAAGAGAGA GTTCTAGGG GCTACTGTT ATGAGACACA TCCAGAGAT ATGTTTAAGT 2160
AAAAATCCTT GAGAAATTAT TATGTCAGAT GTTTTTCAT TCATTATCAG GAAGTTTATG 2220

```

5
 10
 15

```

TTATCTGTCA TTTTITTTIT TCACATCAGT TTGATCAGGA AAGTGTATAA CACATCTTAG 2280
AGCAAGAGTT AGTTTGGTAT TAAATCCTCA TTAGAACAAC CACCTGTTTC ACTAATAACT 2340
TACCCCTGAT GAGTCTATCT AAACATATGC ATTTAAGCC TTCAATATAC ATTATCAACA 2400
TGAGAGAAAT AACCAACAAA GAAGATGTTC AAAATAATAG TCCCATATCT GTAATCATAT 2460
CTACATGCAA TGTAGTAAAT TCTGAAGTIT TTTAAATTIA TGGCTATTTT TACACGATGA 2520
TGAATTTTGA CAGTTTGTGC ATTTTCTTTA TACATTTTAT ATCTTCTGT TAAAATATCT 2580
CTTCAGATGA AACTGTCCAG ATTAATTAGG AAAAGGCATA TATTAACATA AAAATTGCAA 2640
AAGAAATGTC GCTGTAAATA AGATTACAA CTGATGTTTC TAGAAAATTT CCACTTCTAT 2700
ATCTAGGCTT TGTCACTAAT TTCCACACCT TAATTATCAT TCAACTTGCA AAGAGACAAA 2760
CTGATAAGAA GAAAATTGAA ATGAGAATCT GTGGATAAGT GTTTGTGTTT AGAAGATGTT 2820
GTTTGTCCAG TATTAGAAAA TACTGTGAGC CGGCATGCT GGTTCATATC TGTAATCCCA 2880
GCACTTTGGG AGGCTGAGGG GGTGGATCAC CTGAGGTCGG GAGTTCTAGA CCAGCCTGAC 2940
CAACATGGAG AAACCCATC TCTACTAAAA ATACAAAATT AGCTGGGCAT GGTGGCACAT 3000
GCTGGTAATC TCACTATTG AGGAGGCTGA GGCAGAGAAA TTGCTTGAAC CCGGAGGCG 3060
GAGGTTGCAG TGAGCCAAAG TTGACCACT GTACTCCAGC CTGGGTGACA AAGTCAGACT 3120
CCATCTCCAA AAAAAAAAAA AAAA
  
```

Seq ID NO: 20 Protein sequence
 Protein Accession #: NP_055146.1

20
 25
 30

```

1 11 21 31 41 51
| | | | |
MVRKEPVSTI SKGGYLGQNV NGRLPGLGNK EPPGQEKVQL KRKVTLLRGV SIIIGTIIGA 60
GIFISPKQVL QNTGSGVMSL TIWTVGVLS LFGALSYAEL GTTIKK9GQH YTYILEVFGP 120
LPAFVVRWVE ILIIRPAATA VISLAFGRYI LEFFFIQCEI PELAIKLITA VGITVVMVLN 180
SMSVSWASRI QIFLTFCKLT AILIIIVPGV MQLIKGQTON FKDAPSGRDS SITRLPLAFY 240
YGMAYAGGTF YLNFVTBEVE NPEKTIPLAI CISMALITGV YVLTNVAYFT TINAEELLVS 300
NAVANTFSEK LLGNFSLAVP IFVALSCFPG MNGGVFAVSR LFYVASREGE LPEILSMIHV 360
RKHTPLPAVI VLRLPTMIML FSGDLSLLN FLSPARWLEI GLAVAGLIYL RYKCPDMHIRP 420
PKVPLFIPAL FSTCTLPVMA LSLYSDPFST GIGFVITLIG VPAYYLFILN DKKPRWFRIM 480
SEKITRTILQI ILEVVPEDSK L
  
```

Seq ID NO: 21 DNA sequence
 Nucleic Acid Accession #: NM_002422.2
 Coding sequence: 64..1497

40
 45
 50
 55
 60
 65
 70

```

1 11 21 31 41 51
| | | | |
ACAAGGAGGC AGGCAAGACA GCAGGCGATA GAGACAACAT AGAGCTAAGT AAAGCCAGTG 60
GAAATGAAGA GCTCTCCAAT CCTACTGTTG CTGTGCGTGG CAGTTTGCTC AGCCTATCCA 120
TTGGATGGAG CTGCAAGGGG TGAGGACACC AGCATGAACC TTGTTAGAAA ATATCTAGAA 180
AACTACTAGC ACCTCAAAA AGATGTGAAA CAGTTTGTTA GGAGAAAGGA CAGTGGTCTT 240
GTTGTTAAAA AAATCGAGAG AATGCAGAG TTCTTGGAT TGGAGGTGAC GGGGAAGCTG 300
GACTCCGACA CTCTGGAGGT GATGCGCAAG CCCAGGTGTG GAGTTCCTGA TGTGTGTCAC 360
TTCAAGACCT TCTCTGGCAT CCGAAGTGG AGGAAAACCC ACCTTACATA CAGGATTGTG 420
AATATACAC CAGATTGCC AAAAGATGCT GTTGATTCIG CTGTTAGAAA AGCTCTGAAA 480
GTCCTGGGAG AGGTGACTCC ACTCACATTC TCCAGGCTGT ATGAAGGAGA GGCTGATATA 540
ATGATCTCTT TTGCACTTAG AGAACATGGA GACTTTTACC CTTTGTAGTG ACCTGGAAAT 600
GTTTGTGCCC ATGCTATGTC CCTTGGGCCA GGGATTAATG GAGATGCCCA CTTTGTATGAT 660
GATGAACAAAT GGACAAAGGA TACAACAGGG ACCAATTTAT TTCTCTTGC TGTCTATGAA 720
ATGCGCCACT CCCTGGGTCT CTTTCACTCA GCCAACACTG AAGCTTTGAT GTACCCACTC 780
TATCACTCAC TACTAGAACT GACTCGGTTC CGCTGTCTC AAGATGATAT AAATGGCACT 840
CAGTCCCTCT ATGGAACCTCC CCTGACTCC CCGTGAACCC CCTGTGTACC CACGGAACCT 900
GTCCCTCCAG AGCCTGGGAC GCCAGCCAACT TGTGATCCTG CTTTGTCTCT TGAATCTGTC 960
AGCACTCTGA GAGGTGAAAT CCTGATCTTT AAGACAGGCG ACTTTTGGGG CAATCCCTC 1020
AGGAAGCTTG AACCTGAATT GCATTTGATC TCCTCATTTT GGCCATCTCT TCCTTCAGGC 1080
GTGGATGCGG CATATGCAAG GACCTGTTT TCAATTTTAA AGGAAATCAA 1140
TTCTGGGCCA TCAGAGGAAA TGAGGTACGA GCTGGATACC CAAGAGGCAT CCACACCTTA 1200
GGTTTCCCTC CAACCTGAG GAAATCGAT GCAGCCATTT CTGATAGGA AAGAACAAA 1260
ACATATTTCT TTGTAGAGGA CAAATACTGG AGATTGTATG AGAAGAGAAA TTCCATGGAG 1320
CCAGGCTTTC CCAAGCAAAT AGCTGAAGAC TTTCCAGGGA TTGACTCAA GATTGATGCT 1380
GTTTGTGAAG AATTGGGTTT CTTTATTTC TTTACTGGAT CTTCACAGTT GGAGTTTGAC 1440
CCAAATGCAA AGAAAGTGAC ACACACTTTG AAGAGTAACA GCTGGCTTAA TTGTTGAAAG 1500
AGATATGTAG AAGGCACAAT ATGGGCACCT TAAATGAAGC TAATAATTCT TCACCTAAGT 1560
CTCTGTGAAT TGAATGTTT GTTTCTCTCT GCCTGTGCTG TGACTCGAGT CACACTCAG 1620
GGAACTTGAG CGTGAATCTG TATCTTCCCG GTCAATTTTA TGTATTTACA GGGCATTCAA 1680
ATGGGCTGCT GCTTAGCTTG CACCTTGTCA CATAGAGTGA TCTTTCCCAA GAGAAGGGGA 1740
AGCACTCGTG TGCAACAGAC AAGTGACTGT ATCTGTGTAG ACTATTGCT TATTTAATAA 1800
AGACGATTG TCAGTGTGT T
  
```

Seq ID NO: 22 Protein sequence
 Protein Accession #: NP_002413

75
 80

```

1 11 21 31 41 51
| | | | |
MKSLEPILLL CVAVCSAYPL DGAARGEDTS MNLVQKYLEN YDLEKDVQK FVERKDSGFP 60
VKKIREMQKF LGLEVTKGLD SDTLEVMRKP RGVEDVGHF RFPGIPKWR KTHLYTRIVN 120
YTFDLPKDAV DSAVEKALKV WEEVTLPLTS RLYEGBADIM ISFAVREBD FYFDDPGPNV 180
LAHAYAPGPG INGDAHFDD EQWTKDTTGT NLFLVAHEI GHSLGLPESA NYBALMYFLY 240
ESLDTLRFER LSDQDINGIQ SLYGPPFDSE EPLVFTPEFV PPEPGTPANC DPALSFDAVS 300
TLRGEILFVR DRHFWKSLR KLEPELHLIS SFWPSLPSGV DAAYEVTSKD LVFIFKGNQF 360
WAIRGNVEKA GYPRGIETLG FPTVVRKIDA AISDKENKNT YFFVEDKYWR FDEKRNMSMF 420
GFPKQIAEDF FGIDSKIDAV PEEFGFFYFF TGSSQLEFDP NAKKVTHTLK SNBWLNC
  
```

Seq ID NO: 23 DNA sequence
Nucleic Acid Accession #: NM_006528
Coding sequence: 57..764

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

```

1      11      21      31      41      51
|      |      |      |      |      |
GCCGCCAGCG GCTTCTCTGG ACGCCCTTGCC CAGCGGGGCG CCGGACCCCC TGCAACCATGG 60
ACCCCGCTCG CCCCTTGGGG CTGTCTGATTC TGCTGCTTTT CTGACCGGAG GCTGCACTGG 120
GCGATGCTGC TCAGGAGCCA ACAGGAAATA ACGCGGAGAT CTGTCTCTCG CCCCTAGACT 180
ACGACCCCTG CCGGGCCCTA CTTCTCGGTT ACTACTACGA CAGGTACACG CAGAGCTGCC 240
GCCAGTTTCT GTACGGGGGC TCGAGAGGCA ACGCCAACAA TTTCTACACC TGGGAGGCTT 300
GCGACGATGC TTGCTGGAGG ATAGAAAAG TTCCCAAAGT TTGCGGCTG CAAGTGAGTG 360
TGGACGACCA GTGTGAGGGG TCCACAGAAA AGTATTCTTT TAATCTAAGT TCCATGACAT 420
GTGAAAATTT CTTTTCGGGT GGGTGTCAAC GGAACCGGAT TGAGAACAGG TTTCAGATG 480
AAGCTACTTG TATGGGCTTC TGCGCACCAA AGAAATTTCC ATCATTTTGC TACAGTCCAA 540
AAGATGAGGG ACTGTGCTCT GCCAATGTGA CTCGCTATTA TTTTAAATCCA AGATACAGAA 600
CCTGTGATGC TTTCACCTAT ACTGCTGTG GAGGGAATGA CAATAACTTT GTTAGCAGGG 660
AGGATTGCAA ACGTGCATGT GCAAAAGCTT TGAAGAAAG AAAGAAGATG CCAAAGCTTC 720
GCTTTGCCAG TAGAATCCGG AAAATTCCGA AGAAGCAATT TTAACATTC TTAATATGTC 780
ATCTTGTGTT TCTTTATGTC TTATTGCTCT TTAGGTGTGT ATCTGAAGA TAATATGACA 840
GCATGAGGAA ACAAAATCAT GGTGATTAT TCACCACTTT TTATTAATAC AAGTCACTTT 900
TTCAAAATTT TGAATTTTT TATATATAAC TAGCTGCTAT TCAAAATGTA GTCTACCAAT 960
TTTAAATTTT GGTTCACACTG TTTGTGAGAG GAATCTCTGC AATGCATAAG ATATAAAAGC 1020
AAATATGACT CACTCATTTT TTGGGCTCGT ATTCTGATT TCAGAAAGAG ATCATAACTG 1080
AAACAACATA AGACAATATA ATCATGTGCT TTTAACATAT TTGAGAATAA AAGGACTAG 1140
CC
  
```

Seq ID NO: 24 Protein sequence
Protein Accession #: NP_006519

30
35
40
45
50
55
60
65
70
75
80

```

1      11      21      31      41      51
|      |      |      |      |      |
MDPARPLGLS ILLLEFLTEAA LGDAAQEPTG NNAEICLLPL DYBPCRALLL RYYDRTYQS 60
CRQFLYGGCE GWANNFYTWE ACDDACWRIE KVPKVCRLQV SVDDQCEGST EKYFFALSSM 120
TCEKFFPGGC EHRRIENRFP DEATCMGFCA PKKIPSPCY S PKDEGLCSAN VTRYVFNPRY 180
RTCDAFYTYG CQGNNDNFVS REDCKRIACAK ALIKKKKKMPK LRFASRIKI RKKQF
  
```

Seq ID NO: 25 DNA sequence
Nucleic Acid Accession #: NM_005458.1
Coding sequence: 1..2825

40
45
50
55
60
65
70
75
80

```

1      11      21      31      41      51
|      |      |      |      |      |
ATGGCTTCCC CGGGAGGCTC CGGGCAGCCA GGGGGGCGGC CGCGCGCCGC ACCGCGCCCC 60
GCGCGCCTGC TACTGCTACT GCTGCTGCGG CTGCTGCTGC CTCTGGCGCC CCGGGCCTGG 120
GGCTGGGCGC GGGGGCGCCC CCGGCGCGCG CCCAGCAGCC CCGCGCTCTC CATCATGGGC 180
CTCATGCGCG TCACCAAGGA GGTGGCCAG GGCAGCATCG GCGCGGTGT GCTCCCGGCC 240
GTGGAACCTG CCATCGAGCA GATCCGCAAC GAGTCACTCC TGCGCCCTCA CTCTCTGAC 300
CTGCGGCTCT ATGACACGGA GTGCGACAAC GCAAAAGGGT TGAAGAGCTT CFACGATGCA 360
ATAAAATACG GCGCGAACCA CTGTATGTTG TTTGGAGGCG TCTGTCCATC CGTCACATCC 420
ATCATTTGCG AGTCCCTCCA AGGCTGGAAT CTGGTGCGAG TTTCTTTTGC TGCAACCAAG 480
CCTGTTCTAG CCGATAAGAA AAAATACCTT TATTTCTTTC GGACCGTCCC ATCAGACAAT 540
GCGGTGAATC CAGCCATTTT GAGGTGTGTC AAGCACTACC AGTGGAGAG CGTGGGCAAG 600
CTGACGCAAG ACCTTCAGAG GTTCTCTGAG GTGCGGAATG ACCTGACTGG AGTTCTGTAT 660
GGCGAGGACA TTGAGATTTC AGACACCGAG AGCTTCTCCA ACATCCCTG TACCACTGTC 720
AAAAAGCTGA AGGGGAATGA TGTGCGGATC ATCCTTGCC AGTTTGACCA GAATATGGCA 780
GCAAAAGTGT TCTGTTGTGC ATACGAGGAG AACATGATG GTAGTAAATA TCAGTGGATC 840
ATTTCGGGCT GGTADGAGCC TTCTTGTGTG GAGCAGGTGC ACACGGAGC CAATCATCC 900
CGCTGCTTCC GGAAGAATCT GCTTGTGCTC ATGAGGGGCT ACATTTGCGT GGATTTGAG 960
CCCTGAGCT CCAAGCAGAT CAAGACCATC TCAGGAAAGA CTCACAGCA GTATGAGAGA 1020
GAGTACAGCA ACAGCGGTC AGGCGTGGG CCCAGCAAGT TCCACGGTGA CGCTACGAT 1080
GGCATCTGGG TCATGSCCAA GACACTGCAG AGGGCCATGG AGACACTGCA TGCCAGCAGC 1140
CGGACCAAGC GATCCAGGA CTTCAACTAC ACAGACCTCA CGCTGGGCG GATCATCCTC 1200
AATGCCATGA ACCGAGACCA CTCTCTGGG GTCAAGGGTC AAGTTGTATT CCGGAATGGG 1260
GAGAGAATGG GGAACATTAA ATTACTCAA TTCAAGACA GCAGGGAGGT GAGGTGGGA 1320
GAGTACAAAG CTCTGGCCGA CACACTGGAG ATCATCAATG ACACCATCAG GTTCCAAGGA 1380
TCCGAACCA CAAAGACAA GACCATCATC CTGGAGCAGC TGCGGAAGAT CTCCCTACCT 1440
CTCTACAGCA TCCCTCTGCG CCTCACCATC CTGGGATGTA TCATGGCCAG TGCCTTTCTC 1500
TTCTTCAACA TCAAGAACCG GAATCAGAA GTCATAAAGA TGTGAGTCC ATACATGAAC 1560
AACCTTATCA TCCCTGAGG GATGCTCTCC TATGCTTCCA TATTTCTCTT TGGCCTGAT 1620
GGATCTTTG TCTCTGAAAA GACCTTTGAA ACACCTTGCA CCGTCAAGAC CTGGAATCTC 1680
ACCGTGGGCT ACACGACCGC TTTTGGGGCC ATGTTTGCAA AGACCTGGAG AGTCCACGCC 1740
ATCTTCAAAA ATGTGAAAAA GAAGAGAGAG ATCATCAAG ACCGAAACT GCTTGTGATC 1800
GTGGGGGCGA TGTGCTGATG GAGCTGTGAT ATCTGATCT GCTGGCAGG TGTGGACCCC 1860
CTGCGAAGGA CAGTGGAGAA GTACAGCATG GAGCGGAGC CAGCAGGAGC GATATCTCC 1920
ATCGCCCTCT TCTGTGAGCA CTGTGAGAAC ACCCATATGA CCATCTGGCT TGGCATGCTC 1980
TATGCTTACA AGGCACTTCT CATGTTGTTT GGTGTTTTCT TAGCTTGGGA GACCCGCAAC 2040
GTACAGATCC CGCACTCAA GCACAGCAAG TACATGGGGA TGAATGTCTA CAACTGGGG 2100
ATCATGTGCA TCATCGGGG CGCTGTCTCC TTCTGAGCC GGGACCAACC CAATGTGCA 2160
TTCTGATCAG TGGCTCTGCT CATCATCTTC TGACAGCACA TCACCTCTG CCGTGTATT 2220
GTGCGAAGC TCATCAACCT GAGAACAAAC CCAGATGCAG CAACGCGAA CAGGCGATTC 2280
CAGTTCATCT AGAATCAGAA GAAAGAGAT TCTAAAACCT CCACCTCGGT CACCACTGTC 2340
AACCAAGCCA GCATATCCCG CTTGGAGGGC CTACAGTCAG AAAACCATCG CCTGCAATG 2400
  
```

5
AAGATCACAG AGCTGGATAA AGACTTGGAA GAGGTACCA TCAGCTGCA GGACACACCA 2450
GAAAGACCA CCTACATTAA ACAGAACCCAC TACCAAGAGC TCAATGACAT CCTCAACCTG 2520
GGAACCTTCA CTGAGAGCAC AGATGGAGGA AAGGCCATTT TAAAAATCA CCTCGATCAA 2580
AATCCCCAGC TACAGTGGAA CACAACAGAG CCTCTCGAA CATGCAAGA TCCTATAGAA 2640
GATATAAACT CTCCAGAACCA CATCCAGCGT CGGCTGTCC TCCAGCTCCC CATCTCCAC 2700
CAGCCTTACC TCCATCCAT CGGAGGCGTG GACCCAGCT GTGTGAGCCC CTGCTCAGC 2760
CCCAACGCCA GCCCCGCCA CAGACATGTG CCACCTCTCT TCCGAGTCAT GGTCTCGGC 2820
CTGTAA

10 Seq ID NO: 26 Protein sequence
Protein Accession #: NP_005449.1

15
1 MASRRSGQP GRPPPPPPPP ARLLLLLLLP LLLPLAGAW GNARGAPRPP PSSEPLSIMG 60
IMPLTKEVAK GSIGRGVLPA VELAEQIRN ESLLRPYFLD LRLYDTECDN AKGLKAFYDA 120
IKYGFNHLMV FGGVCPSTVS IIAESLQWNN LVQLSFAATT FVLADKKKYP YFFRTVPSDN 180
AVNPAILKLL KHYQWKEVGT LTQDVQRFSE VRNDLTGVLY GEDIEISDTE SPSNDPCTSV 240
20 KKLKGNVDRI ILGQFDQNMV AKVFOCAYES NMYGSKYQWI IFGNVYEPNWW EQVHTEANSS 300
RCLRNILLAA MEGYTGVDPE PLSSKQIKTI SGKTPQYER EYNNKRSVGK PSKFHGYAYD 360
GIWVIKTLQ RAMEILHASS RHORIQDFNY TDHTLGRIIL NAMNETNFFG VTGQVVRNNG 420
ERMGTIKFTQ FQDSREVKVG EYNAVADTLE IINDTIRFQG SEPPKDKTII LEQLRKISLP 480
LYSLSALTI LGMIMASAPL PFNKNRNQK LKMSSEPMN NLIILGGMLS YASIFLPLGD 540
25 GSPVSEKTFE TLCTVTETWL TVGTTAFGA MFAKTWRVHA IFKNVGMKKK IIXDQKLLVI 600
VGGMLLDLC ILICQAVDP LERTVEKYSM SPDPAGRDIS IRPLLEHCEN THMTIWLIV 660
YAYRGLMLF GCFLAWETRN VSPALNDSK YICMSVYNVG IMCIIGAAVS FLTRDQPNVQ 720
PCIVALVIF CSTITLCLVF VFKLITLRN EDRAATQNRFF QFTONQKKED SKTSTSVTSV 780
NQASTSRLEG LQSENHRLRM KITELDKDLE EVTMQLQDTF EKTYYIKQNE YQELMDILNL 840
30 GNFTSTDDG KALLKMLDQ NFQLQWNTTE PSRTCKDPIB DINSPEHIQR ELSLQLPLIH 900
HAYLPSIGGV DASCVSPCVS PTASPRHRHV PPSFRVNVSG L

35 Seq ID NO: 27 DNA sequence
Nucleic Acid Accession #: NM_000450.1
Coding sequence: 117..1949

35
1 CCTGAGACAG AGGCAGCAGT GATACCCACC TGAGAGATCC TGIGTTTGAA CAACGCTTC 60
CCAAAACGGA AAGTATTTCA AGCCTAAACC TTTGGGTGAA AAGAACTCTT GAAATCATGA 120
40 TTGCTTACA GTTCTCTTCA GCTCTCACC TTGGTCTTCT CATTAAGAGC AGTGGAGCCT 180
GGTCTTACA CACCTCCACG GAAGCTATGA CTTATGATGA GGCCAGTGCT TATTGTACGC 240
AAAGGTACAC ACACCTGGTT GCAATTCAAA ACAAGAAGA GATTGAGTAC CTAAACTCCA 300
TATTGAGCTA TTACACCAAGT TATTACTGGA TTGGAATCAG AAAAGTCAAC AATGTGTGGG 360
45 TCTGGGTAGG AACCCAGAAA CCTCTGACAG AAGAAGCCAA GAACTGGGCT CCAGGTGAAC 420
CCACAATLAG GCAAAAGAT GAGGACTGCG TGGAGATCTA CATCAAGAGA GAAAAGATG 480
TGGGCATGTG GAATGATGAG AGGTGCAGCA AGAAGAAGCT TGCCCTATGC TACACAGCTG 540
CCTGTACCAA TACATCTGTC AGTGGCCACG GTGAATGTGT AGAGACCATC AATAATTACA 600
CTTGCAAGTG TGACCTCTGC TTCAGTGGAC TCAAGTGTGA GCAAAATGTG AACTGTACAG 660
50 CCTGGAATC CCTGAGCAT GGAAGCCTGG TTTCAGTCA CCACTGGGA AACTTCAGCT 720
ACAAATCTTC GTTCTCTATC AGCTGTGATA GGGGTACCT GCCAAGCAGC ATGGAGACCA 780
TGCAATGTAT GTCTCTGGA GAATGGAGTG CTCTATTCC AGCTGCAAT GTGGTTGAGT 840
GTGATGCTGT GACAAATCCA GCCAATGGGT TCGTGAATG TTTCCAAAC CCTGGAAGCT 900
TCCATGAGTG CACAACCTGT ACATTTGACT GTGAAGAAGG ATTTGAACCTA ATGGGAGCCC 960
55 AGAGCCTTCA GTGTACCTCA TCTGGGAAT GGGACAACGA GAAGCCAAAG TGTAAAGCTG 1020
TGACATGAG GGCCTGCGC CAGCCTCAGA ATGGCTCTGT GAGGTGCAGC CATTCCTCTG 1080
CTGGAGATT CACCTTCAA TCATCTGCA ACTTCACTG TGAGGAAGGC TTCAATGTGC 1140
AGGGACCAGT CCAAGTTGAA TGCAACCTC AAGGGCAGTG GACACAGCAA ATCCCAGTTT 1200
GTGAAGCTT CCAGTGCACA GCTTGTCCA ACCCGAGCG AGGCTACATG AATTGTCTTC 1260
60 CTAGTCTTC TGGCAGTTT CATTATGGGT CCAGCTGTGA GTTCTCTGT GAGCAGGGTT 1320
TTGTGTTGAA GGAATCCAAA AGGCTCCAAT GTGGCCCCAC AAGGGAGTGG GACAACGAGA 1380
AGCCCACTG TGAAGCTGTG AGATGCAATG CTCTCCACA GCCCCGAGG GGTTTGGTGA 1440
GGTGTCTCA TTCCCTATT GGAGAATCA CCTACAAGTC CTCTGTGTC TTCACTGTG 1500
AGGAGGGATT TGAATTATAT GGAATCAACT AACTTGAGTG CACATCTCAG GGACAATGGA 1560
65 CAGAAGAGT TCCTTCTGTC CAAGTGTGTA AATGTTCAAG CCTGGCAGTT CCGGGAAGA 1620
TCAACATGAG CTGAGTGGG GAGCCCTGT TTGGCACTGT GTGCAAGTTC GCCTGTCTG 1680
AAGGATGGAC GCTCAATGGC TCTCAGCTC GGACATGTGG AGCCACAGGA CACTGGTCTG 1740
GCCTGCTACC TACCTGTGAA GCTCCCACTG AGTCCAACAT TCCCTTGGA GCTGGACTTT 1800
CTGCTGCTGG ACTCTGCCYC CTGACATTAG CACCATTTCT CCTCTGGCTT CGGAAATGCT 1860
70 TACGAAAGC AAGAAATTT GTTCTTCCA CGACCTGCA AAGCCTTGAA TCAGACGGAA 1920
GCTACCAAAA GCTTCTTAC ATCCTTAAAG TTCAAAAGAA TCAGAAACAG GTGCATCTGG 1980
GGAACTAGAG GGATACACTG AAGTTAACAG AGACAGATAA CTCTCTCGG GTCTCTGGCC 2040
CTCTTGCCCT ACTATGCCAG ATGCCCTTAT GGCTGAAACC GCAACACCCA TCACCCTTC 2100
AATAGATCAA AGTCAGCAG GCAAGGACGG CTTTCAACTG AAAAGACTCA GTGTTCCCTT 2160
75 TCTTACTCTC AGGATCAAGA AAGTGTGGC TAATGAAGGG AAAGGATATT TTCTTCCAG 2220
CAAGGTGAA GAGACAGAGA CTCTGAATC TCAGAAATTC TTTTCTAAT CTCTCTGCT 2280
CGCTGTAAA TCTTGGCACA GAAACAAAT ATTTGTGGC TTTCTTCTT TTGCCCTTCA 2340
CAGTGTTCG ACAGCTGAT ACACAGTTC TGTCTAAGA ATGAATATA ATTATCCAGA 2400
GTTAGAGGA AAAAAATGAC TAAAAATATT ATAACCTAAA AAAATGACAG ATGTTGAATG 2460
80 CCAACAGGCA AATGCAATGA GGGTGTGTA TGGTGCAAAT CCTACTGAAT GCTCTGTGG 2520
AGGGTACTA TGCACAATTT AATCACTTTC ATCCCTATGG GATTCACTGC TTCTTAAAGA 2580
GTCTTAAAG ATTGTGATAT TTTTACTTGC ATTGAATATA TATTAATCTT CCACTTCTT 2640
TCATTCAATA CAAGTGTGGT AGGACCTTAA AAACTTGTG AATGCTGTCA ACTATGATAT 2700
GGTAAAGTT ACTTATCTTA GATTACCCCC TCATTGTTTA TTAACAAATT ATGTTACATC 2760
TGTTTAAAT TTAATTCAAA AAGGAAACT ATTTGCCCT AGCAAGGCAT GATGTTAACC 2820

5
10
15

```

AGAATAAAGT TCTGAGTGT TTTACTACAG TTGTTTTTGG AAAACATGGT AGAATTGGAG 2880
AGTAAAAAAT GAATGGGAAGG TTTGTATATT GTCCAGATATT TTTTCAGAAA TATGTGGTIT 2940
CCACGATGAA AACTTTCAT GAGGCCAAAC GTTTTGAAC TATAAAGCA TAAATGCATA 3000
CACACAAAGG TATATTTTA TGAATGTCCT TGTGGAAAA GAATACAGAA AGATGGATGT 3060
GCTTTGCATT CCTACAAGA TGTTTGTGAG ATGTGATATG TAAACATAAT TCTGTATAT 3120
TATGGAGAGT TTTAAATCA CARTAGAAAC TCACCATGTA AAGAGTCAT CTGGTAGATT 3180
TTTACGAAT GAAGATGCT AATAGTTATT CCTATTGTG TTTCTTCTGT ATGTTAGGGT 3240
GCTCTGGAAG AGAGGAATGC CTGTGTGAGC AAGCATTTAT GTTTATTAT AAGCAGATT 3300
AACAATCCA AAGGAATCTC CAGTTTTGAG TTGATCAGTG GCAATGAAAA ATTCTCAGTC 3360
AGTAATTGCC AAGCTGCTC TAGCCTTGAG GAGTGTGAGA ATCAAACTC TCCTACACTT 3420
CCATTAACTT AGCATGTGTT GAAAAAATA GTTTCAGAGA AGTCTGGCT GAACACTGGC 3480
AACACAAAG CCAACAGTCA AATCAGAGAT GTGATAAGGA TCAGAACAGC AGAGGTTCTT 3540
TTAAAGGGGC AGAAAACTC TGGGAATAA GAGAGAACAA CTACTGTGAT CAGGCTATGT 3600
ATGGAATACA GTGTATTTT CTTTGAATT GTTAAAGTGT TGTAAATATT TATGTAACT 3660
GCATTGAAA TTAGCTGTGT GAAATACCAG TGTGTTTGT GTTTGAGTT TATGTAGAA 3720
TTTAAATTAT AACTTAAAT ATTTTATAAT TTTTAAAGTA TATATTATT TAAGCTTAG 3780
TCAGACCTAT TTGACATAAC ACTATAAAGG TTGACAATA ATGTGCTTAT GTTT

```

Seq ID NO: 28 Protein sequence
Protein Accession #: NP_000441.1

20
25
30
35

```

1 11 21 31 41 51
| | | | |
MIASQFLSAL TLVLLIKESG ANSYNTSTEA MTYDASATC QQRNYTHLVAI QNKKEIEYLN 60
SILSYSPSY WIGRKVMNV WVVVGTQKPL TEEAKMWAPG EENNRQKDED CVEIYIKREK 120
DVGMMNDERC SKKKLALCYT AACTNTSCSG HGEVETINN YTCKDQPGFS GLKCEQIVNC 180
TALESPEHGS LVCSHPLGNF SYNSSCSISC DRGYLPSSME TMQCMSSGSEW SAPIFACNVV 240
ECDAVINBAN GFVECFQNPQ SPPWNTTCTF DCEEGFELMG AQSLQCTSSG NWDNEKPTCK 300
AVTCAVRPGR QNGSVRCHS PAGEFTFKSS CNFTCEBGF M LQGPQVVECT TQGWTOQIP 360
VCEAFQCTAL EMERGVMNC LPSASGSERY GSSCEFSCEQ GFVLKSKRL QCGCTGENDN 420
EKPTCEAVRC DAVHQPPKGL VRCASPIGE FTYKSSCAPS CSEGFELYGS TOLECTSQGQ 480
WTEEVPSQV VKCSLAVPG KINMSCSGEP VFGTVCKFAC PEGWTLNGSA ARTCGATGHW 540
SGLLPTCEAP TESNIPLVAG LSAAGLSLLT LAPFLMLRK CLRKAKKFVP ASSCQSLESD 600
GSYQKPSYIL

```

Seq ID NO: 29 DNA sequence
Nucleic Acid Accession #: NM_007036
Coding sequence: 56..610

40
45
50
55
60
65
70
75

```

1 11 21 31 41 51
| | | | |
CTTCCCACCA GCAAGAACCA CGACTGGAGA GCGGAGCGGG AGGCAGCTGG GAACATGAA 60
GAGCGTCTTG CTGCTGACCA CGCTCCTCGT GCTGTCACAC CTGGTGGCCG CCTGGAGCAA 120
TAATTATGCG GTGGACTGCC CTCAACACTG TGACAGCAGT GAGTGCAAAA GCAGCCCGCG 180
CTGCAAGAGG ACAGTGTCTG ACGACTGTGG CTGCTGCGGA GTGTGCGCTG CAGGCGCGGG 240
AGAACTTTCG TACCGCACAG TCTCAGGCAT GGAITGCCATG AAGTGTGGCC CGGGGCTGAG 300
GTGTCAAGCT TCTAATGGGG AGGATCCCTT TGGTGAAGAG TTGTGTATCT GCAAGAGCTG 360
TCCCTACGGC ACCTTCGGGA TGGATTGCGAG AGAGACCTGC AACTGCCAGT CAGGCATCTG 420
TGACAGGGGG ACGGGAAAAA GCGTGAAATT CCGCTCTCTC CAATATTCAG TAACCAAGTC 480
TTCCAACAGA TTGTCTTCTC TCACGGAGCA TGACATGGCA TCTGGAGATG GCAATATTGT 540
GAGAGAAAGG GTGTGTAAGG AGAATGCTGC CGGCTCTCCC GTAATGAGGA AATGGTTAAA 600
TCAACGCTGA TCCTGGCTGT GATTTCCTGAG AGAAGGCTCT ATTTTCTGTA TTGTTCACAA 660
CACAGCCAAC ATTTTGGGAA CTTCTAGAT ATAGCATAAG TACATGTAAT TTTTGAAGAT 720
CCAATTTGCG ATGCTGCTG GATCCAGAAA ACAAAAGTA GGATACTTAC AATCCATAAC 780
ATCCATATGA CTGAACACTG GTATGTGTTT GTTAAATATT CGAATGCATG TAGATTGTGT 840
AAATGTGTGT GTATAGTAAC ACTGAAGAAC TAAAAATGCA ATTTAGGTAA TCTTACATGG 900
AGACAGGTCA ACCAAGAGG GAGCTAGGCA AAGCTGAGGA CCGCAGTGAG TCATAATTAGT 960
TCTTTGACTT TGTGTACAT TAATGTTGGG ATATGGAAATG AAGACTTAAG AGCAGAGAA 1020
GATGGGGAGG GGGTGGGAGT GGGAAATAAA ATATTTAGOC CTTCCTTGGT AGGTAGCTTC 1080
TCTAGAATTI AATTGTGCTT TTTTPTTTT TTTGGCTTTG GBAAGAGTCA AATAAACA 1140
ACCGAAAAC CCCTGAAGGA AGTAAGATGT TTGAAGCTTA TGGAAATTTG AGTAACAAAC 1200
AGCTTTGAAC TGAGAGCAAT TTCAAAAGGC TGCTGATGTA GTTCCCGGGT TACCTGTATC 1260
TGAAGGACGG TTCTGGGGCA TAGGAAACAC ATACACTTCC ATAAATAGCT TTAACGTATG 1320
CCACCTCAGA GATAAATCTA AGAAGTATT TACCCACTGG TGGTTTGTGT GTGTATGAAG 1380
GTAATATTAT ATATATTTT ATAAATAAAT GTGTTAGTGC AAGTCATCTT CCTTACCCAT 1440
ATTATATCAT CTCTTGAGGA AAGAAATCTA GTATTATTGT TTGAAATATG TTAGAATAAA 1500
AACCTATGAC TCFATAAGGT TTTCAAACAT CTGAGGCATG ATAAATTTAT TATCCATAAT 1560
TATAGGAGTC ACTCTGGATT TCAAAAATG TCAAAAATG AGCAACAGAG GGACCTTATT 1620
TAAACATAAG TGCTGTGACT TCGGTGAATT TTCATTAA GGTATGAAA TAAGTTTTTA 1680
GGAGTTTGT AAAAGAGAA TCAATTTTCA GCAGAAACA TGTCACTTT AAAATATAGG 1740
TGGAAATAGG AGTATATTTG AAAGAATCTT AGCACAACA GGACTGTTGT ACTAGATGTT 1800
CTTAGGAAT ATCTCAGAAG TATTTTATT GAAGTGAAGA ACTTATTTRA GAATTATTTC 1860
AGTATTAC TGTATTATT TCTTGAAGTT GCGCAACAGA GTTGTGAATG TGTGTGGAAG 1920
GCTTTGAAT GTAAAGCTGC ATAAGCTGTT AGGTTTGTGT TTAAGAGAC ATGTTTATTA 1980
TTGTTCAATA AAAAGAACAA AGATAC

```

Seq ID NO: 30 Protein sequence
Protein Accession #: NP_008967.1

80

```

1 11 21 31 41 51
| | | | |
MKSVLLLTLL LVPALVAAV SNMYAVDCPQ HCDSSBECKSS PRCKRTVLDD CGCCRVCAAG 60
RGETCYRTVS GMDGKCGPG LRCQPSNGED PFGEFPGICK DCFYGTFGMD CRKTCNCSG 120
ICDRGTGKCL KPPFFQYSVT KSSNRVSLT EEDMASGDGN LVREEVVKEW AAGSFVMRW 180

```


LNPR

Seq ID NO: 31 DNA sequence

Nucleic Acid Accession #: NM_000963

Coding sequence: 135..1949

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

```

1      11      21      31      41      51
|      |      |      |      |      |
CAATTGTCAT ACGACTTGCA GTGAGCGTCA GGAGCAGCTC CAGGAACCTC TCAGCAGGCG 60
CTCCCTTCAGC TCCACAGCCA GACGCCCTCA GACAGCAAAG CCTACCCCGG CGCCGCGCCC 120
TGCCCGCCCGC TCGATGCTC GCGCGCGCCC TGCTGCTGTG CGCGGTCTCG GCGCTCAGCC 180
ATACAGCAAA TCCTTGCTGT TCCACCCCAT GTCAAAACCG AGGTGTATGT ATGAGTGTGG 240
GATTTGACCA GTATAAGTGC GATTGTACCC GGACAGGATT CTATGGAGAA AACTGTCTCA 300
CACCGGAATT TTGACAAAG ATRAAATTAT TTCTGAAACC CACTCCAAAC ACAGTGCAC 360
ACATACCTTAC CCACCTTCAAG GGATTTTTGA ACGTTGTGAA TAACATTCCC TTCCTTGGAA 420
ATGCAATTAT GAGTTATGTC TTGACATCCA GATCACATT GATTGACAGT CCACCAACTT 480
ACAATGCTGA CTATGGCTAC AAAAGCTGGG AAGCCTTCTC TAACCTCTCC TATTATACTA 540
GAGCCCTTCC TCCGTGCTCT GATGATTGCC CGACTCCCTT GGSTGTCAAA GGTAAAAAGC 600
AGCTTCTCTGA TCAAAATGAG ATTGTGGAAA AATTGCTTCT AAGAAGAAAG TTCTCCCTG 660
ATCCCCAGGG CTCAAACTATG ATGTTTGCA TCTTTGCCCA GCACCTCAGC CATCAGTTT 720
TCAAGACAGA TCATAGCTGA GGGCCAGCTT TCACCAACGG GCTGGGCCAT GGGGTGGACT 780
TAAATCATAT TTACGGTGAA ACTCTGGCTA GACAGCGTAA ACTGCGCCTT TTCAGGAGT 840
GAAAATGAA ATATCAGATA ATTGATGGAG AGATGTATCC TCCCACAGTC AAAGATACTC 900
AGGCAGAGAT GATTCACCTT CCTCAAGTCC CTGAGCATCT ACGGTTTGCT GTGGGGCAGG 960
AGGTCTTTGG TCTGGTGGCT GGTCTGATGA TGTATGCCAC AATCTGGCTG CGGGAACACA 1020
ACAGAGTATG GAGTTGTGCT AAACAGGAGC ATCTGAATG GGGTATGAG CAGTTGTTC 1080
AGACAAGCAG GCTAATACTG ATAGAGAGA CTATTAGAT TGTGATTGAA GATTATGTGC 1140
AACACTTGAG TGGCTATCAC TTCAAACCTGA AATTGACCC AGAAGTACTT TTCACAAAC 1200
AATTCCAGTA CCAAATGCT ATTGCTGCTG AATTAAACAC CCTCTATCAC TGGCATCCC 1260
TTCTGCTCTG CACCTTTCAA ATTCATGACC AGAAATACAA CTATCAACAG TTTATCTACA 1320
ACAACTCTAT ATTGCTGGAA CATGGAATTA CCAAGTTGTG TGAATCATTC ACCAGGCAAA 1380
TTCTGGCAG GGTGCTGCTG GGTAGGAATG TTCCACCCGC AGTACAGAAA GTATCACAGG 1440
CTTCCATTGA CCAGAGCAGG CAGATGAAAT ACCAGTCTTT TAATGAGTAC CGCAAAACGCT 1500
TTATGCTGAA GCGCTATGAA TCATTGGAAG AACTTACAGG AGAAAAGGAA ATGCTGTCAG 1560
AGTTGGAAGC ACTCTATGCT GACATCGATG CTGTGGAGCT GTATCCTGCC CTTCCTGGTAT 1620
AAAAGCCTCG GCCAGATGCC ATCTTTGGTG AAACCATGGT AGAAGTTGGA GCACCATCTC 1680
CCTTGAAAGG ACTTATGGGT ATGTTATAT GTTCTCTGTC CTACTGGAGG CCAAGCACTT 1740
TTGGTGGAGA AGTGGGTTTT CAAATCTCA ACACCTGCTC AATTCAGTCT CTCATCTGCA 1800
ATAACGTGAA GGGCTGTCCC TTACTTCTAT TCAGTGTTC AGATCCAGAG CTCATTAAAA 1860
CAGTCACCAT CAATGCAAGT TCTTCCCGCT CCGGACTAGA TGATATCAAT CCCACAGTAC 1920
TACTAAAGA ACCTGCTGCT GAACGTAGA AGTCTAATGA TCATATTTAT TTATTATAT 1980
GAACCATGTC TATTAAATTA ATTATTTAAT AATATTTATA TTAACCTCCT TATGTTACTT 2040
AACATCTTCT GTAACAGAG TCAGTACTCC TGTTCGCGAG AAAGGAGTCA TACTTGTGAA 2100
GACTTTTATG TCACCTACTC AAAGATTTTG CTGTGTCTGT TAAGTTTGA AAACAGTTTT 2160
TATTTCTGTT TATAAACCAAG AGAGAAATGA GTTTTGCCTT CTTTTTACTT GAATTTCAAC 2220
TTATATATA AGAACGAAAG TAAAGATGTT TGAATACTTA AACACTATCA CAAAGTGGCA 2280
AAATGCTGAA AGTTTTTACA CTGTGATGT TTCCAAATGA TCTTCCATGA TGCATTAGAA 2340
GTAACATAAG TTTGAAATTT TAAAGTACTT TTGGTTATTT TTCTGTCTAT AAACAAAAAC 2400
AGGTATCAGT GCATTATTTA ATGAAATATT AATTAGACA TTACAGTAA TTTCTATGCT 2460
ACTTTTAAAA ATCAGCAAGT AAACAAATAT TTGAAATTTT TAAATTCATA GGGTAGAATC 2520
ACCTGTAAGA ACCTGTTTGA TTCTTAAAG TTATTAACT TGTACATAA CCAAAAAGAA 2580
GCTGTCTTGG ATTTAAATCT GTAAATCAG ATGAAATTTT ACTACAAATG CTGTGTTAAA 2640
TATTTTATAA GTGATGTTCC TTTTTCACCA AGAGTATAAA CCTTTTATG GTGACTGTTA 2700
AAACTTCTCT TTAATCAAAA ATGCCAATTT TATTAAAGTG GTGGAGCCAC TGCAGTGTTA 2760
TCTCAAAATA AGAATATTTT GTTGAGATAT TCCAGATTTT GTTTATATG CTGGTAACAT 2820
GTAAATCTTA TATCAGCAAA AGGGTCTACC TTTAAATAAA GCAATAACAA AGAAGAAAAAC 2880
CAAATATATG TTCAAATTTA GGTTTAAACT TTTGAAGCAA ACTTTTTTTT ATCCTTGTGC 2940
ACTGCAGGCC TGTACTCAG ATTTTCTAT GAGGTTAATG AAGTACCAG CTGTGCTTGA 3000
ATAACGATAT GTTTTCTCAG ATTTTCTGTT GTACAGTTTA ATTTAGCAGT CCATATCACA 3060
TTGCAAAAGT AGCAATGACC TCATAAATA CCTCTTCAA ATGCTTAAAT TCATTTCACA 3120
CATTAATTTT ATCTCAGTCT TGAAGCCAAT TCAGTAGGTG CATTGGAATC AAGCCTGGCT 3180
ACCTGCATGC TGTTCCTTTT CTTTTCTTCT TTTAGCCATT TTGCTAAGAG ACACAGTCTT 3240
CTCATCACTT CTTTTCTCCT ATTTTGTTTT ACTAGTTTFA AGATCAGAGT TCACTTTCTT 3300
TGGACTCTGC CTATATTTTC TTACCTGAAC TTTTGCAAGT TTTCAGGTAA ACCTCAGCTC 3360
AGGACTGCTA TTAGCTCCTT CTTAAGAAGA TTAAGAAGA AAAAAAAGG CCCTTTTAAA 3420
AATAGTATAC ACTTATTTTA AGTGAAAAGC AGAGAATTTT ATTTATAGCT AATTTTAGCT 3480
ATCTGTAAAC AAGATGGATG CAAAGAGGCT AGTGCTCAG AGAGAAGTGT ACGGGGTTTG 3540
TGACTGGAAA AAGTTACGTT CCCATTCTAA TTAATGCCCT TTCTTATTTA AAAACAAAAC 3600
CAAAATGATAT CTAAGTAGTT CTCAGCAATA ATAATAATGA CGATAATACT TCTTTTCCAC 3660
ATCTCATGT CATGACATTT TAATGGTACT GTATATTACT TAATTTATG AAGATTATTA 3720
TTTATGCTT ATTAGACAC TATGTTATA AACTGTGTTT AAGCCTACAA TCATTGATT 3780
TTTTTTGTTA TGTACAAATC AGTATATTTT CTTGGGGTTT ACCCTCTGTA ATATTATGTA 3840
AACAACTCAA AGAAATGATT GTATTAAGAT TTGTGAATAA ATTTTAGAA ATCTGATTGG 3900
CATATTGAGA TATTTAAGGT TGAATGTTG TCCTTAGGAT AGGCCTATGT GCTAGCCAC 3960
AAAGAAATG GTCTCATTAG CCGTAATGTG CCATAAGACT GACCTTTTAA AATGTTTGA 4020
GGGATCTGTG GATGCTTCGT TAATTTGTTT AGCCACAATT TATTGAGAAA ATATTCTGTG 4080
TCAAGCACTG TGGTTTAA TATTTTAA TCAACGCTG ATTACAGATA ATAGTATTTA 4140
TATAAATAAT TGAATAAAAT TTTCTTTTGG GAAGAGGGAG ABAATGAAAT AATATCATT 4200
AAAGTAACT CAGGAGAACT TTCTTACAA TTTACGTTT AGAATGTTA AGGTTAAGAA 4260
AGAAATAGTC AATATGCTG TATAAACAC TGTTCACGT TTTTTTAA AAAAAAAT 4320
GATTTGTTAT TAACATTGAT CTGCTGACAA AACCCTGGAA TTTGGGTTGT GTATGCGAAT 4380
GTTTCAGTGC CTCAGACAAA TGTGATTTTA ACTTATGTAA AAGATAAGTC TGAATAATAA 4440
TGTCTGTTTA TTTTGTACT ATTTA

```

Seq ID NO: 32 Protein sequence
Protein Accession #: NP_000954

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      MLARALLICA VLALSHTANP CSHPCQNRG VCMVGFEDQY KCDCTRTGFY GENCSTPEFL 60
      TRIKFLFKPT PNTVHYILTH FRGFVNVVMN IPFLRNAMS YVLTSRSHLI DSPPTYNADY 120
      GYKSWFAFSN LSYVTRALPP VPDDCTPLG VKGKKQLPDS NEIVEKLLLR RKFPIDPFQS 180
10     NMMAFFAFQH FTHQFFKTDH KRGAFTNGL GHGVDLNHIY SEILARQKRL RLFKDGKMKY 240
      QIIDGMYPP TVKDTQAEMI YPOVPEHLR FAVQGEVGL VPGLMMYATI WLRHNRVCD 300
      VLKQSEBPMG DBQLFQTSRL ILIGETIKIV IEDYVQHLRG YHFKLKFDEE LLENKQFQYQ 360
      NRIAAEFNTL YHWEPLLEPD FQIHQKINY QQFIYNNIL LEHGTTQFVE SPTRQIAGRV 420
      AGGRNVPPAV QKVSQASIDQ SRQMKYQSFN EYRKRFLKP YSFBEELTGE KMSAELEAL 480
      YGDIADAVELY PALLVEKPRP DAIFGETMVE VGAPFSLKGL MGNVICSPAY WKPSTFGGEV 540
15     GPOINTASI QSLICNVKVG CPFTSFVSPD PELIKTVTIN ASSRSGLDD INPTVLLKER 600
      STEL

```

Seq ID NO: 33 DNA sequence
Nucleic Acid Accession #: NM_001508.1
Coding sequence: 1..1361

```

20     1      11      21      31      41      51
      |      |      |      |      |      |
      ATGGCTTCAC CCAGCCTCCC GGGCAGTGAC TGCTCCAAA TCATGATCA CAGTCATGTC 60
25     CCGAGTTTG AGGTGGCCAC CTGGATCAAA ATCACCCTTA TTCGGTGTA CCGATCATC 120
      TTCTGATGG GCTTCTCGGG GAACAGCGTC ACCATTGGG TCACCCAGST GCTGCAGAG 180
      AAAGGATATC TGCAAGAAGA GGTGACAGAC CACATGGTGA GTTGGCTTG CTGGACATC 240
      TTGTTTTC TCATCGGAT GCGCATGGAG TTCTACAGCA TCATCTGGAA TCCCTGACC 300
      ACCTCCAGCT ACACCTCTGC CTGCAAGCTG CACACTTTC TCTCGAGGC CTGCAGCTAC 360
30     GCTACCTGC TGCACTGCT GAOGCTCAGC TTTGAGCGCT ACATCGCCAT CTGTACCCOC 420
      TTCAGTACA AGCTGTGTC GGGACCTTGC CAGGTGAAGC TGCTGATTGG CTTCGTCTG 480
      GTCACTCCG CCTTGGTGGC ACTGCCCTTG CTGTTGCCA TGGGTACTGA GTACCCCTG 540
      GTGAACGTGC CCAGCCACCG GGTCTCCTCT TGCAACCGCT CCAGCACCCG CCACCCAGAG 600
      CAGCCCGAGA CTTCAATAT GTCCATCTGT ACCAACCTCT CCAGCCCGTG GACCGTGTTC 660
35     CAGTCCAGCA TCTCTCGGC CTTCGTGGT TACCTCGTG TCTGCTCTC CGTAGCCTTC 720
      ATGTGCTGGA ACATGATGCA GGTGCTCATG AAAAGCCAGA AGGCTCTCT GGCCTGGGGC 780
      ACGCGGCTC CCGAGCTGAG GAAGTCCGAG AGCGAAGAGA GCAGGACCGC CAGGAGGCAG 840
      ACCATCATCT TCTGAGGCT GATTGTTGTG ACATTGCGCG TATGCTGGAT GCCCAACAG 900
40     ATTGCGAGGA TCATGCTGC GGCACAAACC AAGCAGCACT GGACGAGGTC CTACTTCCG 960
      GGTACATGA TCTCTCTCC CTCTCGGAG ACGTTTCTT ACCTCAGCTC GGTATCAAC 1020
      CGCTCTCTGT ACACGCTGTC CTGCGAGCAG TTTCGGCGGG TGTGCTGCA GGTGCTGTG 1080
      TGCCGCTGT CGCTGCAGCA CGCCAAACC GAGAAGCGCC TGGCGGTACA TGGCAGCTCC 1140
      ACCAOCGACA GCGCCCGCTT TGTGCGGCG CCGTTGCTCT TCGGCTCCCG GCGCCAGTCC 1200
45     TCTGCAAGGA GAAGTGAAG GATTTCCTTA AGCACTTTC AGAGCGAGGC CGAGCCCCAG 1260
      TCTAAGTCCC AGTCATTGAG TCTGAGTCA CTAGAGCCCA ACTCAGGCCG GAACCCAGCC 1320
      AATTCTGCTG CAGAGAATGG TTTTCAGGAG CATGAAGTTT GA

```

Seq ID NO: 34 Protein sequence
Protein Accession #: NP_001499.1

```

50     1      11      21      31      41      51
      |      |      |      |      |      |
      MASPSLPQSD CSQIIDHSHV PEFEVATWIK ITLILVYLII FVMGLLGNV TIRVTQVLQK 60
      KGYLQKSVTD HMYSLACSDI LVFLIGMPME FYSIWNELT TSSYTLSCKL HTFLFEACSY 120
55     ATLLHVLTLF PERYIALCHP FRYKAVSGPC QVKLLIGFVW VTBALVALPL LFAMGTREYL 180
      VNVPSHGLT CNRSSTRHHE QPTEHMSIC TNLSSRWTFV QSHFGAFVW YLVVLLSVAF 240
      MCWMMQVLM KQXGSLAGG TRPPQLRKE SESSRTARRQ TIIFLRLIVV TLAVCHMFWQ 300
      IRRIMAAKP KHDWTRSYFR AYMLLPFSR TFFYLSVIN PLXYTVSSQ FRRVVFQVLC 360
60     CRLSLQRAHE EKRLRVAHS TTDSARFVQR PLLFASRRQS SARRTKIFL STFQSEABPQ 420
      SKSQSLSLBS LEPNSGAKPA NSAAENGFQK HEV

```

Seq ID NO: 35 DNA sequence
Nucleic Acid Accession #: NM_006475.1
Coding sequence: 28..2538

```

65     1      11      21      31      41      51
      |      |      |      |      |      |
      AACAGACTG CAACGGAGAG ACTCAAGATG ATTCCCTTTT TACCAATGTT TTCTCTACTA 60
70     TTGCTGCTTA TTGTTAACC TATAAACGCC AACATCATT ATGACAAGAT CTGGCTCAT 120
      AGTCGATCA GGGTTCGGGA CCAAGGCCCA AATGTCGTG CCTTCAACA GATTTTGGGC 180
      ACCAAAAGA AATACTTCAG CACTTGTAG AACTGTGATA AAAAGTCCAT CTGTGACAG 240
      AAACGACTG TTTTATATGA ATGTTGCCCT GGTATATGA GAATGGAAG AATGAAGGC 300
      TGCCGACGAG TTTTGGCCAT TGACCATGTT TATGGCACTC TGGGCATCGT GGGAGCCACC 360
75     ACACGCGAGC GCTATTCTGA CGCTCAAAA CTGAGGGAGG AGATCGAGGG AAAGGGATCC 420
      TTCACTTACT TTGCACCGAG TAATGAGGCT TGGGACAACT TGGATTCTGA TATCCGTAGA 480
      GGTTTGGAGA GCAACGTGAA TGTGAATTA CTGAATGCTT TACATAGTCA CATGATTAT 540
      AAGAGAATGT GCAACAAGGA CTTAAAAAAT GGCATGATTA TTCCTTCAAT GTATAACAAT 600
      TTGGGCTTT TCATTAACCA TTATCTTAAT GGGGTGTGCA CTGTTAATTG TGCTCGAATC 660
80     ATCCATGGGA ACCAGATTGC AACCAATGGT GTTCTCCATG TCATTGACCG TGTGCTTACA 720
      CAATTTGGTA CCTCAATTCA AGACTTCATT GAAGCAGAG ATGACCTTTC ATCTTTTAGA 780
      GCAGCTGCCA TCACATCGGA CATATTGGAG GCCCTTGGAA GAGACGGTCA CTTCACACTC 840
      TTGCTTCCCA CCAATGAGGC TTTTGAGAAA CTTCCACGAG GTGTCTTAGA AAGGTTTATG 900
      GGAGACAAAG TGGCTTCCGA AGCTCTTATG AAGTACCACA TCTTAATATC TCTCCAGTGT 960
      TCTGAGTCTA TTAATGGGAG AGCAGTCTTT GAGACGCTGG AAGGAAATAC AATTGAGATA 1020

```

5
10
15
20
25
30
35
40

```

GGATGTGACG GTGACAGTAT AACAGTAAAT GGAATCAAAA TGGTGAACAA AAAGGATATT 1080
GTGACAAATA ATGGTGTGAT CCATTGTGAT GATCAGGTCC TAATTCCTGA TTCTGCCAAA 1140
CAAGTTATTG AGCTGGGCTGG AARACAGCAA ACCACCTTCA CGGATCTTGT GGCCCAATTA 1200
GGCTTGGCAT CTGCTCTGAG GCCAGATGGA GAATACACTT TGCTGGCACC TGTGAATAAT 1260
GCATTTTCGT ATGATACTCT CAGCATGGTT CAGCGCCTCC TTAAATTAAT TCTGCAGAAT 1320
CACATATTGA AAGTAAAGT TGGCCTTAAT GAGCTTTACA ACGGGCAAAT ACTGGAAACC 1380
ATCGGAGGCA AACAGCTCAG AGTCTTCGTA TATCGTACAG CTGTCTGCAT TGAATAATCA 1440
TGCAATGGAG AAGGGAGTAA GCAAGGGAGA AACGGTGGA TTACATATT COGOGAGATC 1500
ATCAAGCCAG CAGAGAAATC CCTCCATGAA AAGTTAAAC AAGATAAGCG CTTTAGCACC 1560
TTCCCTAGCC TACTTGAAGC TGCAGACTTG AAGAGCTCC TGACACAACC TGGAGACTGG 1620
ACATTATTTG TCCCAACCAA TGTATGCTTT AAGGGAATGA CTAGTGAGGA AAAAGAAATT 1680
CTGATACGGG ACACAAATGC TCTTCAAAAC ATCATTTCTT ATCACCCTGAC ACCAGGAGTT 1740
TTCAATGGAA AAGGATTGTA ACCTGGTGT ACTAACATT TAAAGACCAC ACAAGGAAGC 1800
AAAATCTTTC TGAAGAAGT AAATGATACA CTTCGTGTGA ATGAATTGAA ATCAAAAGAA 1860
TCTGACATCA TGACACAAA TGGTGTAAAT CATGTTGTAG ATAACTCCT CTATCCAGCA 1920
GACACACCTG TTGGAATGTA TCAACTGCTG GAAATACCTA ATAAATTAAT CAATACATC 1980
CAAAATTAAG TTGTTCTGTG TAGCACCTTC AAGAAATCC CGTGACTGT CTATACAACT 2040
AAAATTATAA CCAGAAGTTGT GGAACCAAAA ATTAAAGTGA TTGAAGGCAG TCTTCAGCCT 2100
ATTATCAAAA CTGAAGGACC CACACTAACA AAGTCAAAA TTGAAGGTGA ACCTGAATTC 2160
AGACTGATTA AAGAAGGTGA AACAATAACT GAAGTGATCC ATGGAGAGCC AATTATTAAA 2220
AAATACACCA AAATCATTTGA TGGAGTGCCT GTGGAATTA CTGAAAAGA GACACGAGAA 2280
GAACGAATCA TACAGGCTCC TGAATAAAAA TACACTAGGA TTCTACTGG AGGTGGAGAA 2340
ACAGAAGAAA CTCTGAAGAA ATTGTTACAA GAACAGGTCA CCAAGGTGAC CAAATTCTAT 2400
GAAGGTGGTG ATGCTCATTT ATTTGAAGAT GAAGAAATTA AAGAGCTGCT TCAGGGAGAC 2460
ACACCGTGCA TGACAGTTCG AGCCAACAAA AAGTTCAGG GTTCTAGAG ACGATTAAAG 2520
GAAGGTGCTT CTCAGTGAAG ATCCAAAAAC CAGAAAAAAA TGTATTATCA ACCCTAAGTC 2580
AATAACCTGA CCTTAGAATA TTGTGAGAGC CAAGTTGACT TCAGGAACCTG AAACATCAGC 2640
ACAAAAGAAA AATCATCAAA TAATTCAGAA CACAAATTTA ATATTTTTTT TCTGAATGA 2700
GAAACATGAG GGAATTTGTG GAGTTAGCCT CCTGTGTGTA AGGAATTGAA GAAATATATA 2760
CACCTTACAC CTCTTTTCAT CTGACATTA AAGTCTCTGG CTAACCTTGG AATCCATTAG 2820
AGAAAAATCC TTGTCAACCG ATTCATTACA ATTCAAATCG AAGAGTTGTG AACTGTTATC 2880
CCATTGAAA GACCGAGCCT TGTATGTATG TTATGGATAC ATAAATGCA CGCAAGCCAT 2940
TATCTCTCCA TGGGAAGCTA AGTTATAAAA ATAGGTGCTT GGTGTACAAA ACTTTTTATA 3000
TCAAAAGGCT TTGCACATTT CIATATGAGT GGGTTACTG GTAAATTATG TTATTTTTTA 3060
CAACTAATTT TGTACTCTCA GAATGTTTGT CATATGCTTC TTGCAATGCA TATTTTTTAA 3120
TCTCAAAAGT TTCAATAAAA CCATTTTTCA GATATAAAGA GAATTACTTC AATTGTAGTA 3180
ATTCAAAAA ACTCAAGATT TAAGTTAAAA AGTGGTTTGG ACTTGGGAA

```

Seq ID NO: 36 Protein sequence
Protein Accession #: NP_006466.1

45
50
55
60

```

1 11 21 31 41 51
| | | | |
MIPFLPMFSL LLLLVNFIN ANNHDKILA HSRIRGRDQG PNVCALQQL GKKKYPSTC 60
KNWYKKSICG QCTTVLYECC PGYMRMEGMR GCPAVLPIDH VYGTGLGVGA TTTQRYSDAS 120
KLRESIEGKG SFTYFAPSNE AWDNLDSDIR RGLSNVNVVE LLMLHSEMI NKRMLTKDLK 180
NGMILIPSMY NLGLFINHYE NGVVTVNCAR IIHGNQIATN GVHVHIDRVL TQIGTSIQDF 240
IEAEDDLSSP RAAAITSDIL EALGRDGHFT LFAPTNEAFE KLPGRVLERF MGDKVASEAL 300
MKVEILLNTLQ CSBSIMGGAV FETLENTIE IGCDGDSITV NGIKMVKKKD IVTNNGVIHL 360
IDQVLIPDSA QVVISLAKGQ QTFPDLVAQ LGLABALRPD GEYTLAPVN NAFSDDTLSM 420
VGRLLKLILQ NHILKVKVGL NELYNGQILE TIGGKQLRVF VYRTAVCIEN SCMEKGSKQG 480
RNGAIHIFRE IHPAKSKSLH EKLQDKRFS TFLSLLEAAD LKELLTQPGD WTLFVPTNDA 540
EKGMTSEKES ILRLDKNALQ NILLYHLTFG VPIGKGFEFG VTNLLKTQGS SKIFLKEVND 600
TLVNLKSLKQ ESDIMTNGV IHVVDKLLYP ADTPVGNQDL LELINKLIK YIQKPVVRSST 660
FKELPVTVTY TKIITKVVPEP KLVIEGSLQ PIKTEGPTL TKVKIEGEPE FRLIKEGETI 720
TSVIEGEPII KKYTKIIDGV FVEITEKETR BERIITGPBI KYTRISTGGG ETEETLKKLL 780
QSEVTKVTKF IEQGDGHLFE DEEIKRLQG DTPVKLQAN KKVQSSRRRL REGRSQ

```

Seq ID NO: 37 DNA sequence
Nucleic Acid Accession #: NM_002416
Coding sequence: 40..417

65
70
75
80

```

1 11 21 31 41 51
| | | | |
ATCCAATACA GGAGTGACTT GGAACCTCCAT TCTATCACTA TGAAGAAAAG TGGTGTTCIT 60
TTCCCTCTGG GCATCATCTT GCTGGTCTCG ATTGGAGTGC AAGGAACCCC AGTAGTGAGA 120
AAGGGTGGCT GTTCTCTGAT CAGCACCAAC CAAGGGACTA TCCAACCTACA ATCCTTGAAA 180
GACCTTAAAC AATTGTCCCC AAGCCCTTCC TCGAGAAAAA TTGAAATCAT TGCTACACTG 240
AAGAAATGAG TTCAAAACATG TCTAAACCCA GATTGAGCAG ATGTGAAGGA ACTGATTAAA 300
AAGTGGGAGA AACAGGTCAG CCAAAAGAAA AAGCAAAAGA ATGGGAAAAA ACATCAAAAA 360
AAGAAAGTTC TGAAGATTCT AAAATCTCAA CGTTCTCGTC AAAAGAAGAC TACATAAGAG 420
ACCACTTCAC CAATAAGTAT TCTGTGTTAA AAATGTTCTA TTITAATTAT ACGGCTATCA 480
TTCCAAAGGA GGATGGGATA TAATACAAAG GCTTATTAAAT TTGACTAGAA AATTTAAGAA 540
ATTACTCTGA AATTGTAATC AAAGTTAGAA AGTTGATTTT AAGAATCCAA ACCTTAAGAA 600
TGTGTAAGG CTATGATGT CTGTGTTCTT CTACCACCCA CCAAGTTGAA TTCAATATGC 660
TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT GTTCAACCAA CCACATCCCA 720
CTCACAACAG CTGCGCTGGA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCAGAGAG 780
TATCTGAGGC ACATGTCTAG AAGTCTTAAG CCGTTTAGCA TGCTGGTGAG CCAGCAGTT 840
TGAATTCAG CTGACCTCA CCAAGCTGCT GTGGCATCA ACCTCTGTAT TTGAATCAGC 900
CTCAGGCTAG CACACAAT GTGTCTGAGA GATTATGCT GATTGTTATT GGGTATCACC 960
ACTGGAGATC ACCAGTGTGT GGCCTTCAGA GCCTCCTPTC TGGCTTTGGA ABCCATGTGA 1020
TTCCATCTTG CCCGCTCAGG CTGACCACTT TATTTCTTTT TGTTCGCCCT TGCTTCATTC 1080
AAGTCAGCTC TTCTCCATCC TACCACAATG CAGTGCCCTT CTCTCTTCCA GTGCACCTGT 1140
CATATGCTCT GATTATCTG AGTCAACTCC TTTCTCATCT TGTCCCAAC ACCCCACAGA 1200

```

5 AGTGCCTTCT TCTCCCAATT CATCCTCACT CAGTCCAGCT TAGTTCAGT CCTGCCTCTT 1260
 AAATAAACCT TTTTGGACAC ACRAATTATC TTAATACTCC TGTTCACITT GGTTCAGTAC 1320
 CACATGGGTG AACACTCAAT GGTTAACATA TTTTGGGTG TTTATCCTAT CTCTCCAAAC 1380
 AGATTGTGAG CTCTCTGAGG GCAAGAGCCA CAGTATATTT CCTGTCTTCT TCACACAGTGC 1440
 CTAATAATAC TGTGGAACTA GGTTTTAATA ATTTTAAAT TGATGTGTGT ATGGGCAGGA 1500
 TGGCAACACG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCCTGGCTAC FCCATGTTGG 1560
 CTAGCCTCTG GTAACCTCTT ACTTAITATC TTCAGGACAC TCACTACAGG GACCAAGGAT 1620
 GATGCAACAT CCTGTCTTTT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA 1680
 GCACGTGGTA AAACACTTGC GGATATTTCT GACTGTTTTT AAAAAATATA CAGTTTACCG 1740
 10 AAATCATAT AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC 1800
 CCNACCATAC AAAAATCTCT TTTCCGAGG GAAAAAGGCT TTCTCAATAA GCCTCAGCTT 1860
 TCTAAGATCT AACCAAGATG CCACCGAGAT CCTTATCGAA ACTCATTTTA GGCAAAATAG 1920
 AGTTTTATTG TCCGTTTACT TGTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA 1980
 TCTCCCATGA AGAAAGGGAA CGGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT 2040
 15 TATGGAAAGC ATGATTGGTG CCCAGTAGC CTCGACAGGA TGTGGAAACC TCCTCCAGG 2100
 GGAGGTTTCA TGAATTGTGT AGGAGAGGTT GTCTGTGGCC AGAATTTAAA CCTTACTCTA 2160
 CTTCCTCCAA TTGAATCACT GCTCACACTG CTGATGATTT AGAGTGTCTGT CCGGTGGAGA 2220
 TCCCAACCGA ACBTGCTTAT TAATCATGAA ACTCCCTAGT TCCTTCATGT AACTTCCCTG 2280
 AAAAACTTAA GTGTTTCATA AATTGAGAG TCTGTGACCC ACTTACCTTG CATCTCACAG 2340
 20 TAGACAGTA TATGAATAAC AACCAAGAC TACATATTGT CACTGACACA CAGTTATATA 2400
 TCATTATACA TATATATACA TACATGCATA CACTCTCAA GCAAATAATT TTTCACTTCA 2460
 AAACAGTATT GACTGTATA CCTGTAAAT TGAAATATT TCITTTGTAA AATAGAATGG 2520
 TATCAATAAA TAGACATTA ATCAG

25 Seq ID NO: 38 Protein sequence
 Protein Accession #: NP_002407

30 1 11 21 31 41 51
 MKKSGVLFLL GIILLVLIGV QGTEVVRGR CSCISTNQGT IHLQSLKDLK QFAPSPSCEK 60
 ILIIATLKNQ VQTCINPDFA DVKELIKKWE KQVSQKQKQK NGKIKHQKKV LKVRKSQRSR 120
 QKRTT

35 Seq ID NO: 39 DNA sequence
 Nucleic Acid Accession #: NM_006670
 Coding sequence: 85..1347

40 1 11 21 31 41 51
 CCGGCTCGCG CCTTCGGGCG CAGCCTCTCC GAGCCCTCGG AGCGGGCGGC GTCCAGCCCC 60
 AGCTCCGGGG AAACCGGAGC CGCGATGCCT GGGGGGTGCT CCCGGGGCCC GCGCCCGGGG 120
 GACGGGGCTC TGGCGCTGGC GCGACTAGCG CTGGTACTCC TGGGCTGGGT CTCTCTGTCT 180
 TCTCCCACTC CTCTGGCATC CTCTCTCTCC TCCTCGGGGC CGTTCTCTGC TTCCGCGGTG 240
 45 TCCGCGGAGC CCGCGCTGCC GGAACGATGC CCGCGCTGTG GCGAGTGTCT CGAGGCGAGC 300
 CGCACAGTCA AGTGGCTTAA CCGCAATCTG ACCGAGGTGC CCAACGACCT GCGGCGCTAC 360
 GTGCGCAACC TCTTCTTAC CCGCAACGAG CTGGCGGTGC TCCCTGCGCG GCGCTTCGCC 420
 CGCGCGCGCG CGCTGGCGGA GCTGGCGCGG CTCAACCTCA GCGGCGAGCG CCTGGACGAG 480
 GTGCGCGCGG GCGGCTTCGA GCATCTGCCG AGCCTGCGCC AGCTCGACCT CAGGCCAAC 540
 50 CCACTGGCGC ACCTCAGTCC CTTCGCTTTC TGGGCGAGCA ATGCCAGCCT CTCGCGCCCC 600
 AGTCCCTCTG TGGAACTGAT CTTGAACACAC ATCGTGCCCC CTGAAGATGA GCGGCGAGAA 660
 CGGAGCTTCG AGGGCATGCT GGTGGCGGCC CTGCTGCGCG GCGGTGCACT GCAGGGGCTC 720
 CGCGCGCTTG AGCTGGCGAG CAACCACTTC CTTTACCTGC CGCGGGATGT GCTGGCCCAA 780
 TCGCCAGGCC TACGTCACCT GCACTTAAGT AATAATTGCG TGGTGAGCCT GACCTACGTG 840
 55 TCCTTCGSCA ACCTGACACA TCTAGAAAGC CTCCACCTGG AGGACAATGC CCTCAAGGTC 900
 CTTCAAGATG GCACCTGGC TGAAGTGCAA GGTCTACCCC ACATTAGGGT TTTCTGGGAC 960
 AACCAATCCT GGTCTGCGA CTGCCACATG GCAGACATGG TGAACCTGCT CAAGGAACA 1020
 GAGGTAGTGC AGGGCAAGA CCGGCTCACC TGTGCATATC CGGAATAAT GAGGAATCGG 1080
 GTCTCTCTGG AGCTGACAG TGTGACCTGC GACTGTGACC CGATTCTTCC CCAATCCCTG 1140
 60 CAACCTCTTT ATGCTTCTCT GGTATTGTT TTAGCCCTGA TAGGGGCTAT TTTCTCTCTG 1200
 GTTTGTATT TGAACCGCAA GGGGATAAAA AAGTGGATGC ATAAACATCG AGATGCGCTG 1260
 AGGGATCACA TGGAGGGGTA TCATTACAGA TATGAAATCA ATGCGGACCC CAGATTAACA 1320
 AACCTCAGTT CTAACTCGGA TGTCTGAGA ATATTAGAGG ACAGACCAAG GACAACTCTG 1380
 CATGAGATGT AGACTTAAGC TTTATCCCTA CTAGGCTTGC TCACATTTCA TCTTCCACTA 1440
 65 TAGATACAA CCACTTTGAC TAAAGCAGT GAAGGGGATT TGCTTCTTGT TTATGTAAAG 1500
 TTTCTCGGTG TGTCTGTGTA ATGTAAGAGC ATGAACAGTT GTGTATAGTG TTTTACCTTC 1560
 TTTCTTTTCT TGGAACTCTT CAACACGTAT GGAGGGGATT TTAGGTTTC AGCATGAACA 1620
 TGGGCTTCTT GCTGTCTGTC TCTCTCTCAG TACAGTTCAA GGTGTAGCAA GTGTACCCAC 1680
 ACAGATAGCA TTCAACAAAA GCTGCTCTCA CTTTTCGAG AAAAAATAC TATTATATAA 1740
 70 TATCAGTTT ATTCTCATGT ACCTAAGTGG TGGAGAAAAA AATTGCATCC TATAAAGTGC 1800
 CTGCAGAGCT TAGCAGGCTC TTCAAATAAA CTCCATGCTG CACAGGAGCA CCTGCATCCA 1860
 AGAGCATGCT TACATTTTAC TGTTCGTCAT ATTCAAAAAA ATAACTTGCA ACTTCATAAC 1920
 TTTCTTGACA AAGTAATAAT CTTTTTGTAT TGCAGTTTAT ATGAAATGT ACTGATTTTT 1980
 75 TTTTAATAAA CTCATCGAG ATCCACCGA CTGAATTGTT AAAAAAATAA AAAAATAAAG 2040
 ATTCTAAAA GAA

80 Seq ID NO: 40 Protein sequence
 Protein Accession #: NP_006661

1 11 21 31 41 51
 MPGGCSRGPA AGDGRLELRL LALVLLGWVS SSSPTSSASS FSSSAPPLAS AVSAQPPLED 60
 QCPALCECSE AARTVKKVNR MLFEVPTDLP AYVRHLFLYG NQLAVLPAGA FARRPPLAEL 120
 AALNLSGSR LPSLRQLDLS HNPPLADLSPF AFSGSNASVS APSPLVLLKL 180
 NHTVPPFEDER QNRSPFGMVV AALLAGRALQ GLRRLRLASN HFLVLPDVL AQLPSLRHL 240

LSNMSLVSLT YVSRNLTSL ESLHLEDNAL KVLHNGTLAE LQGLPHIRVF LDNNPWVDCD 300
 RMADMVTKLK ETEVYQKDR LTCAYPEKMR NRVLLELSA DLDGDPILPP SLQTSYVFLG 360
 IVLALIGAIF LLVLYLNRKG IKMMHNRIRD ACRDHMEGYH YRYEINADPR LTNLSGNSDV

Seq ID NO: 41 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..927

1 11 21 31 41 51
 | | | | |
 ATGCTCTGGG GGTGCTCCCG GGGCCCGGCC GCGGGGACG GCGGTCTGCG GCTGGGCGGA 60
 CTAGCCTGGG TACTCTCTGG CTGGGTCTCC TCGTCTTCTC CCACCTCCTC GGCATCCTCC 120
 TTCTCTCTCT CCGCGCGGTT CTGGGCTTCC GCGGTGTCCG CCCAGCCCCC GCTGCCGGAC 180
 CAGTGCCCGC CGCTGTGCGA GTGCTCCGAG GCAGCGCGCA CAGTCAAGTG CGTTAACCGC 240
 AATCTGAOCG AGGTGCGCCG GACCTGCGCC GCTTAOCGTG GCAACCTCTT CCTTACCGGC 300
 AACCACTGGG CCGACAACCA CTCTCTTTAC CTGCGCGCGG ATGTGCTGGC CCAACTGCCC 360
 AGCCTCAGCG ACCTGGACTT AAGTAATAAT TCGCTGGTGA GCCTGACCTA CGTGTCTCTC 420
 CGCAACCTGA CACATCTAGA AAGCTCTCAC CTGGAGGACA ATGCCCTCAA GGTCTCTCAC 480
 AATGGCACCC TGGCTGAGTT GCAAGGTCTA CCCCACATTA GGGTTTTCCT GGACAAACAT 540
 CCGTGGGTCT GCGACTGCCA CATGGCAGAC ATGGTGACCT GGCCTCAAGG AACAGAGGTA 600
 GTGACGGGCA AAGACCGGCT CACCTGTGCA TATCGGAAA AATGAGGAA TCGGTCTCTC 660
 TTGGAACTCA ACAGTCTGTA CCTGGACTGT GACCCGATTC TTCCCCCATC CCGCAAAACC 720
 TCTTATGTCT TCGTCTGATG TGTCTTAGCC CTGATAGGCG CTATTTTCCT CCGTCTTTTG 780
 TATTTGAACC GCAAGGGGAT AAAAAAGTGG ATGCATAACA TCAGAGATGC CTGACGGGAT 840
 CACATGGAAG GGTATCATTA CAGATATGAA ATCAATGCGG ACCCCAGATT AACAAACCTC 900
 AGTTCTAATC CGATGTCTCT CGAGTGA

Seq ID NO: 42 Protein sequence
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 | | | | |
 MGGGCSRGPA AGDGRLLRLAR LALVLLGWVS SSSPTSSASS FSSSAPFLAS AVSAQFFLPD 60
 QCPALCECSE AARTVKCVNR NLTEVPTDLP AYVRLFLTG NQLASNHFLY LPRDVLQALP 120
 SLRHLDLNN SLVSLTVVSP RNLTHLSLE LEDNALKVLH NGTLAELQGL PHIRVFLDNN 180
 FVVCCHMAD MVTNLKETEV VQKDRITCA YPEKMRNRVL LELNSADLDC DPILPPSLOT 240
 SYVFLGIVLA LIGAFILVLV YLNRKGIKKW MENIRDACRD HMEGYHYEYE INADPRILNL 300
 SSSNDVLE

Seq ID NO: 43 DNA sequence
 Nucleic Acid Accession #: NM_058173
 Coding sequence: 68..340

1 11 21 31 41 51
 | | | | |
 AGCGCCTTGC CTCTCTTAG GCTTTGAAGC ATTTTGTCT GTGCTCCCTG ATCTTCAGGT 60
 CACCAACATG AAGTCTTAG CAGTCTGGT ACTCTGGGA GTTTCATCT TTCTGGTCTC 120
 TGCCCGAAT CCGACAACAG CTGCTCCAGC TGACAGTAT CCAGCTACTG GTCTGTCTGA 180
 TGATGAAGCC CCGTATGCTG AAACCACTGC TGCTGCAACC ACTGCGACCA CTGCTGTCTC 240
 TACCACTGCA ACCACGCTG CTCTACACAC TGCTGTAAA GACATTCAG TTTTACCCAA 300
 ATGGGTGGG GATCTCCGAA ATGTAGAGT GTGCTCTGAA GATGGATCA GCTTGAGTCT 360
 TCTGCAATG GGTCACAACT ATTATGCTT CTGTGATTT CATCCAATA CTTACCTTGC 420
 CTACGATATC CCGTTTATCT CTATCAGTT TATTTTCTTT CAAATAAAAA ATACTATGA 480
 GCGAGCTAAC AT

Seq ID NO: 44 Protein sequence
 Protein Accession #: NP_477521

1 11 21 31 41 51
 | | | | |
 MKFLAVLVLV GVSIFLVSAQ NPTTAAPADT YPATGPADDE APDAETTAAA TTATTAAPTT 60
 ATTAASTTAR KDIPVLPKWV GDLNQRVCF

Seq ID NO: 45 DNA sequence
 Nucleic Acid Accession #: NM_000095.1
 Coding sequence: 26..2299

1 11 21 31 41 51
 | | | | |
 CAGCACCCAG CTCCCCGCA CCGCCATGGT CCGGACACC GCTTGCCTTC TTCTGCTCAC 60
 CCTGGCTGCC CTGCGGCGGT CCGGACAGGG CCAGAGCCCG TTGGGCTCAG ACCTGGGCCC 120
 CGAGATGCTT CCGGAACCTG AGGAACTCAA CGCGGCGCTG CAGGACGTGC GGGACTGGCT 180
 GCGGCGGCGG GTCAGGGAGA TCAGTTCTCT GAAAAACAG GTGATGGAGT GTGACGCGTG 240
 CGGGATGCA CAGTCACTAC GCACCGGCTT ACCCGGCTG CCGCCCTGTC TCCACTGCGC 300
 GCGCGGCTTC TGCTTCCCG GCGTGGCGTG CATCCAGAG GAGAGCGGGG TGCGCTGGGG 360
 CCGCTGCCCC GCGGGCTTCA CCGGCAAGGG CTGCGACTGC ACCGACGTCA ACGAGTGCAG 420
 CGCCCAACCC TGCTTCCCG GAGTCCGCTG TATCAACACC AGCCCGGGGT TCGCTGCGGA 480
 GCGTGGCGCG CCGGGGTACA GCGGCCCCAC CCACCAAGGC GTGGGGCTGG CTTTGGCCAA 540
 GCGCAACAG CAGGTTTGCA CCGACATCAA CAGGTGTGAG ACCGGGCAAC ATAACTGCGT 600
 CCGCAACTCC GTGTGCATCA ACACCGGGG CTCTCTCCAG TGCGGCGCGT GCGAGCCCGG 660
 CTGCTGGGCG GACCAAGCGT CCGGCTGCCA GCGCGGCGCA CAGCGCTTCT GCGCGAGCGG 720
 CTGCGCCAGC GAGTGCACAG AGCATGCAGA CTGGGTCTTA GAGCGCGATG GCTCGCGGTC 780
 GTGCGTGTGT CCGGTGGGCT GCGCGGCGAA CCGGATCTTC TGTGGTGGCG ACGTGCACCT 840
 AGACGGCTTC CCGGACGAGA AGCTGGGCTG CCGGAGCGCG CAGTGCCTGA AGGACAACTG 900

5
10
15
20
25

```

CGTGACTGTG CCCAACTCAG GGCAGGAGGA TGTGGACCGC GATGGCATCG GAGACCCCTG 960
CGATCCCGAT GCCGACGGGG ACAGGGGTCCC CAATGAAAAG GACAACTGCC CGCTGGTGCG 1020
GAACCCAGAG CACGCGCAACA CGGACGAGGA CAAGTGGGGC GATGCGTGCG ACAACTGCCG 1080
GTCCCGAAG AACGACGACC AAAAGGACAC AGACGAGGAC GGCAGGGGGC ATGCGTGCGA 1140
CGACGACATC GACGGCGACC GGATCGGCAA CCAGGCGGAC AACTGCCCTA GGGTACCCAA 1200
CTCAGACCAG AAGGACAGTG ATGGCGATGG TATAGGGGAT GCCTGTGACA ACTGTCCCCA 1260
GAAGAGCAAC CCGGATCAGG CGGATGTGGA CCACGACTTT GTGGGAGATG CTTGTGACAG 1320
CGATCAAGAC CAGGATGGAG ACGGACATCA GGACTCTCGG GACAACTGTC CCACGGTGCC 1380
TAACAGTGCC CAGGAGGACT CAGACCAAGA TGGCCAGGGT GATGCGTGGC ACGACGAGCA 1440
CGACATGAGT GGAGTCCCTG ACAGTCGGGA CAACTGCCGC CTGGTGCTTA ACCCGGGCCA 1500
GGAGGACGCG GACAGGAGAG GCGTGGCGGA CGTGTGCCAG GACGACTTTG ATGCAGACAA 1560
GGTGGTAGAC AAGATCGAGG TGTGTCCGGA GAACGCTGAA GTACGCTCA CCGACTTCAG 1620
GGCCITCCAG ACAGTCGTGC TGGACCCGGA GGGTGAACCG CAGATTGACC CCAACTGGGT 1680
GGTGTCAAC CAGGGAAGGG AGATCGTGCA GACATGAAC AGCGACCCAG GCCTGGCTGT 1740
GGGTACTACT GCCTTCAATG GCGTGGACTT CGAGGCGACG TTCCATGTGA ACACGGTGAC 1800
GGATGACGAC TATGCGGGCT TCATCTTTGG CTACCAGGAC AGCTCCAGCT TCTACGTGGT 1860
CATGTGGGAG CAGATGGAGC AAACGTATTG GCAGGCGAAC CCTTCCGTG CTGTGGCCGA 1920
GCCTGGCATC CAACTCAAGT CTGTGAAGTC TTCCACAGGC CCGGGGGAAC AGCTGGGAA 1980
CGCTCTGTGG CATAAGGAG ACACAGAGTC CCAGGTGCGG CTGCTGTGGA AGGACCCGCG 2040
AAACGTGGGT TGGAGAGACA AGAAGTCCTA TCGTTGGTTC CTGCAGCAAC GGCCTCAAGT 2100
GGGCTACATC AGGGTGGGAT TCTATGAGGG CCCTGAGCTG GTGGCCGACA GCAACGTGGT 2160
CTTGGACACA ACCATGCGGG GTGGCCGCGT GGGGGTCTTC TGCTTCTCCC AGGAGAACAT 2220
CATCTGGGCC AACCTGCGTT ACCGCTGCAA TGACACCATC CCAGAGGACT ATGAGACCCA 2280
TCAGCTGCGG CAAGCCTAGG GACCAAGGTC AGGACCCGCC GGATGACAGC CACCTTCACC 2340
GCGGCTGGAT GGGGCTCTG CACCCAGCCC AAGGGGTGGC CGTCTGAGG GGAAGTGGG 2400
AAGGCTCAG AGAGACAAA ATAAAGTGTG TGTGCAGGG

```

Seq ID NO: 46 Protein sequence
Protein Accession #: NP_00086.1

30
35
40
45

```

1 11 21 31 41 51
MVFDATCVLL LTLAALGASG QGQSPLSSDL GPQMLRELQE TNAALQDVRD WLRQGVREIT 60
FLKNTVMECD ACNQQSVRT GLPSVRLPLH CAPGFCPPGV ACIQTSSGGR CGPCPAGFTG 120
NGSHCTDVNE CMAMPCEFRV RCINTSEGRF CEACPFYSG PTHQGVGLAF AKARKQVCTD 180
INECBTQQRN CVFNSVCINT RGSFQCGPOQ PGFVGDQASG CQRGAQRFCP DGSFSECHER 240
ADCVLERDGS RSCVCRGWA GNGILCGRDT DLDGFFDEKL RCPEPQCRKD NCVTVFNSGQ 300
EDVDRDIGND ACDPDADGDG VPKNDNCPL VRNPDQRNTD EDKNGDACDN CRSQKNDQK 360
DTDQDGGGDA CDDIDJGRI RNQADNCFRV PMSDQKDSBG DGIQDADQNC PQKSNPDQAD 420
VDHDFVGDAC DSDQDQDGDG RQDERDNCPT VNSAQEDSD HDGQGDACDD DDDNDGVFDS 480
KDNCRLVFNE QGEDADRDGV GDVCQDDFDA DKVVDKIDVC PENAEVTLTD FRAEQTVVLD 540
PSGDAQIDPN VVVINQGRBI VQIMNSDPGL AVGYTAFNGV DFBGTFRVNT VTDDYAGFI 600
FOYQSSSFY VVMKQMGQT YWQANPPRAV AEPGIQLKAV KESTGPGBQL RNALMHTGDT 660
ESQVRLWMD PRNVGWKDKK SYRWFLQHRP QVGYIKRVFY BPELVADSN VVLDITMRGG 720
RLGVPCFSQE NIIWANLYR CNDTIPEDYE TEQLRQA

```

Seq ID NO: 47 DNA sequence
Nucleic Acid Accession #: NM_001565.1
Coding sequence: 67..363

50
55
60
65
70

```

1 11 21 31 41 51
GAGACATTCC TCAATGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC TCCAGTCTC 60
AGCACCATGA ATCAAACTGC GATTCTGATT TGCTGCTCTA TCTTTCTGAC TCTAAGTGCC 120
ATTCAAGGAG TACCTCTCTC TAGAACGTA CGCTGTACCT GCATCAGCAT TAGTAATCAA 180
CCTGTATATC CAAAGTCTTT AGAAAACTT GAAATTATTC CTGCAAGCCA ATTTGTGCTA 240
CGTGTGTAGA TCATTGCTAC AATGAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA 300
TCGAAGGCCA TCAAGAAATT ACTGAAAGCA GTTAGCARGG AAATGICTAA AAGATCTCCT 360
TAAACCAGCA GGGAGCAAA ATCGATCGAG TGCTTCCAAG GATGGACCAAC ACAGAGGCTG 420
CCTCTCCCAT CACTTCCCTA CATGGAGTAT ATGTCAAGCC ATAATTGTTT TTAGTTTACA 480
GTTACACTAA AAGGTBACCA ATGATGGTCA CCAATCAGC TGCTACTRCT CCTGTAGGAA 540
GGTTAATGTT CATCATCTTA AGCTATTGAG TAATAACTCT ACCCTGGCAC TATAATGTAA 600
GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTCTGACC CTGCTTCAA TATTTCCTC 660
ACCTTCCCA TCTTCCAAGG GTACTAAGGA ATCTTCTGCT TTTGGGGTTT ATCAGAAITC 720
TCGAATCTCT AAATAACTAA AAGGTATGCA ATCAAACTG CTTTTAAAG AATGCTCTTT 780
ACITCATGGA CTCTCACTGC CATCTCCCA AGGGGCCCAA ATCTTCTCAG TGGCTACCTA 840
CATACAATTC CAAACACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAAATATT 900
CTTATTTAAT GAAAGACTGT ACAAAGTATA AGTCTTAGAT GTATATATTT CCTATATTGT 960
TTTCAGTGTA CATGGAATAA CATGTAATTA AGTACTATGT ATCAATGAGT AACAGGAAAA 1020
TTTTAAAAAT ACAGATAGAT ATATGCTCTG CATGTTACAT AAGATAAATG TGCTGAATGG 1080
TTTTCAAATA AAAATAGGGT ACTCTCTCTG AATATTTAAG

```

Seq ID NO: 48 Protein sequence
Protein Accession #: NP_001556.1

75
80

```

1 11 21 31 41 51
MQTALICC LIFATLGGIQ GVPLSRTVRC TCISISNQPV NPSLEKLEI IPASQFCPRV 60
ELIATMKKKG SKRCLMFESK AIKNLLKAVS KEMSKRSP

```

Seq ID NO: 49 DNA sequence
Nucleic Acid Accession #: XM_057014
Coding sequence: 143..874

	1	11	21	31	41	51	
	GGGAGGGAGA	GAGGCGCGCG	GGTGAAAGGC	GCATTGATGC	AGCCTGCGGC	GGCCTCGGAG	60
5	CGCGCGCGAG	CCAGACGCTG	ACCACGTTCC	TCTCCTCGGT	CTCCCTCCGC	TCCAGCTCCG	120
	CGCTGCCCGG	CAGCCGGGAG	CCATGCGACC	CCAGGGCCCC	GCOCCTCCDC	CGCAGCGGCT	180
	CCGCGGCCCTC	CTGCTGCTCC	TGCTGCTGCA	GCTGCCCGCG	CCGTGAGGCG	CCTCTGAGAT	240
	CCCCAAGGGG	AAGCAAAAGC	CGCAGCTCCG	GCAGAGGGAG	GTGGTGGACC	TGTATATAGG	300
	AATGTGCTTA	CAAGGGCCAG	CAGGAGTGCC	TGGTCGAGAC	GGGAGCCCTG	GGGCCAATGG	360
10	CATTCCGGGT	ACACCTGGGA	TCCCAGGTGG	GGATGGATTC	AAAGGAGAAA	AGGGGGAATG	420
	TCTGAGGGAA	AGCTTTGAGG	AGTCTGGGAC	ACCCAACTAC	AAGCAGTGTT	CATGGAGTTC	480
	ATTGATATTAT	GGCATAGATC	TGGGAAAAT	TGCGAGTGTT	ACATTTACAA	AGATGCGTTC	540
	AAATAGTGCT	CTAAGAGTTT	TGTTCACTGG	CTCACTTCGG	CTAAATGCA	GAATGCAATG	600
	CTGTGAGGCT	TGGTATTTCA	CATTCAATGG	AGCTGAATGT	TCAGGACCTC	TTCCCATTTG	660
15	AGCTATAATT	TATTGGGACC	AAGGAAGCCC	TGAAATGAAT	TCAACAATTA	ATATTCATCG	720
	CACCTCTTCT	GTGGAAGGAC	TTTGTGAAGG	AATTGGTGCT	GGATTAGTGG	ATGTTGCTAT	780
	CTGGGTTGGC	ACTTGTTCAG	ATTACCCAAA	AGGAGATGCT	TCTACTGGAT	GGAAATTCAGT	840
	TTCTCGCATC	ATTATTGAAG	AACTACCAAA	ATAAATGCTT	TAATTTTCAT	TTGCTACCTC	900
	TTTTTTTATT	ATGCCCTTGA	ATGGTTCAC	TAAATGACAT	TTTAAATAAG	TTTATGTATA	960
20	CATTCTGAATG	AAAAGCAAGG	CTAAATATGT	TTACAGACCA	AAGTGTGATT	TCACACTGTT	1020
	TTTAAATCTA	GCATTATCTA	TTTGTCTTCA	ATCAAAAGTG	GTTTCAATAT	TTTTTTTAGT	1080
	TGGTTAGAAAT	ACTTCTCTCA	TAGTCACATT	CTCTCAACCT	ATAATTTGGA	ATATTGTGTG	1140
	GGTCTTTTGT	TTTTTCTCTT	AGTATAGCAT	TTTTAAAAAA	ATATAAAGC	TACCAATCTT	1200
	TGTACAAATT	GTAAATGTTA	AGAATTTTTT	TTATATCTGT	TAAATAAAAA	TTATTTCCAA	1260
25	CAACCTTAAA	AAAAAATAAA	AAAA				

Seq ID NO: 50 Protein sequence
Protein Accession #: XP_057014

	1	11	21	31	41	51	
30	MRPQGPAAAP	QRLRGLLLLL	LLQLPAPESA	SEIPKKGOKA	QLRQREVVDL	YNGMCLQGPA	60
	GVFGRDGPSP	ANGIPGTGPI	PGRDGFKGEK	GECLRESFEE	SWTFNYKQCS	WSSLNLYGIDL	120
	GKIAECTFTK	MRSNSALRVL	FSGSLRLKCR	NACQQRWYFT	PNGABCSGPL	PIEATIIYLDQ	180
35	GSPEMNSTIN	IHRTBSVEGL	CEGIGAGLVD	VAIWVGTCSD	YPKGDASTGW	NEVSRIIIEE	240
	LPM						

Seq ID NO: 51 DNA sequence
Nucleic Acid Accession #: NM_020974
Coding sequence: 81..3080

	1	11	21	31	41	51	
	GGCGTCCGCG	CACACCTCCC	CGCGCCGCGG	CGGCCACCGC	CGGCACTCCG	CGGCTCTGCG	60
45	CCGCAACCGG	TGAGCCATCC	ATGGGCGTCC	CGGGCCGCAA	CGGTCCCGGG	CGGCGCTGGG	120
	CGGTGCTGCT	GCTGCTGCTG	CTGCTGCGCG	CACGCTGCTG	GCTGGCGGGG	GCCTGCCGCG	180
	CGGGTCCGGG	CGGTGCCGCG	GGGCGCGCAG	AGGATGTAGA	TGAGTGTGCC	CAAGGGCTAG	240
	ATGACTGCCA	TGCCGACGCC	CTGTGTGAGA	ACACACCCAC	CTCCTACAAG	TGCTCTTGCA	300
	AGCCTGGCTA	CCAAGGGGAA	GGCAGGCGAG	GTGAGGACAT	CBATGAATGT	GGAAATGAGC	360
50	TCAATGAGAG	CTGTGTCCAT	GACTGTTTGA	ATATTCAGG	CAATTATCGT	TGCACCTGTT	420
	TTGATGGCTT	CATGTTGGCT	CATGACGGTC	ATAATTGTCT	TGATGTGGAC	GAGTGCGCTG	480
	AGAAATATGG	CGGCTGCCAG	CATACCTGTG	TCAACGTGAT	GGGAGGCTAT	GAGTGTGCTG	540
	GCAAGGAGGG	GTTTTTCTCG	AGTGAACAAT	AGCACACCTG	CAITTCACCG	TGGAAGAGGG	600
	GCTTGAGCTG	CGTGAATAAG	GATCAGCGCT	GATGTCACAT	CTCAGAGGAG	GGCCCAAGGG	660
55	GCAGCGTCCG	CTGTGAGTGC	AGGCGCTGTT	TGAGCTGGCC	CAAGAACCAG	AGAGACTGCA	720
	TCTTGACCTG	TAAACCTGGG	AACGGTGGGT	GCCAGCACTC	CTGTGACGAT	ACAGCCGATG	780
	GCCGAGAGTG	CAGCTGCCAT	CCACAGTACA	AGATGCACAC	AGATGGGAGG	AGCTGCCCTG	840
	AGCGAGAGGA	CAGCTGCCAT	GAGGTGACAG	AGAGCAACAC	CACATCAATG	GTGGATGGGG	900
	ATAAAGCGGT	GAAACCGGGG	CTGCTCATGG	AAACGTGTGC	TGTCAACAAT	GGAGGCTGTG	960
60	ACCGCACTCG	TAAAGTACT	TGCACAGGTG	TCCACTGCAG	TTGCTCTGTT	GGATTCATTC	1020
	TCCAGTTGGA	TGGGAAGACA	TGTAAAGATA	TIGATGAGTG	CCAGACCCGC	AATGGAGGTT	1080
	GTGATCATTT	CTCAAAAAC	ATCGTGGGCA	GTTTGTGACTG	CGGCTGCAGG	AAAGGATTTA	1140
	AATATTAAAC	AGATGAGAAG	TCTTGCCAAG	ATGTGGATGA	GTGCTCTTGG	GATAGGACCT	1200
	GTGACCAACG	CTGCATCAAC	CACCCCTGCA	CATTGTGCTG	TGCTTGCAAC	CGAGGGTACA	1260
65	CCCTGTATGG	CTTCACCCAC	TGTGGAGACA	CCAATGAGTG	CAGCATCMAC	AACGGAGGCT	1320
	GTGACAGGCT	CTGTGTGAC	ACAGTGGGCA	GCTATGAAATG	CCAGTGCCAC	CCTGGGTACA	1380
	AGCTCCACTG	GAATAAAAAA	GACTGTGTGG	AAGTGAAGGG	GCTCCGCCCC	ACAAGTGTGT	1440
	CACCCCTGTG	GTCCCTGCAC	TGCGTAAAGA	GTGCTGGAGG	AGACGGGTGC	TTCTCAGAT	1500
	GTCACTCTGG	CATTCACTTC	TCTTCAGATG	TCACCACCAT	CAGGACAGT	GTAACCTTTA	1560
70	AGCTAAATGA	AGGCAAGTGT	AGTTTGAAAA	ATGCTGAGCT	GTTCOCGAG	GGTCTGCGAC	1620
	CAGCACTACC	AGAGAAGCAC	AGCTCAGTAA	AAGAGAGCTT	COGCTACGTA	AACCTTACAT	1680
	GCAGCTCTGG	CAGCAAGTGC	CCAGGAGGCC	CTGGCCGACC	AAGCACCCCT	AAGGAATATG	1740
	TTATCACTGT	TGAGTTTGGG	CTTGAACCTA	ACCHAAAGGA	GCTGACAGCT	TCTGTGTACC	1800
	TGAGCTGCAT	CGTAAAGCGA	ACCGAGAAGC	GGCTCCGTAA	AGCCATCCGC	AGCTCAGAA	1860
75	AGGCGGTCCA	CAGGGAGCAG	TTTCACCTCC	AGCTCTCAGG	CATGAACCTC	GACGTGGCTA	1920
	AAAAGCCTCC	CAGAACATCT	GAAAGCCAGG	CAGAGTCTGT	TGGAGTGGGC	CAGGGTCTAG	1980
	CAGAAAACCA	ATGTGTGAGT	TGCAGGCGTG	GGACCTATTA	TGATGGAGCA	CGAGAACGCT	2040
	GCATTTTATG	TCCAAATGGA	ACCTTCCAAA	ATGAGGAAGG	ACAAATGACT	TGTGAACCAT	2100
	GCCCCAAGACC	AGGAATTTCT	GGGGCCCTGA	AGACCCACAG	AGCTTGGAAAT	ATGCTGTGAT	2160
80	GTGGAGGTCT	GTGTCAACCT	GGTGAATATT	CTGCAGATGG	CTTTGCACCT	TGCCAGCTCT	2220
	GTGCCCTGGG	CAGGTTCCAG	CCTGAAGCTG	CTCGAACTTC	CTGCTTCCCT	TGTGGAGGAG	2280
	GCCTTGCCAC	CAAAATCAG	GGAGCTACTT	CCTTTCAGCA	CTGTGAACCC	AGAGTTCAAT	2340
	GTTCACTTGG	CAATTTCTAC	AACACCAACA	CTCACCGATG	TATTCGTGGC	CCAGTGGGAA	2400
	CATACCAGCC	TGAATTTGGA	AAAAATTAAT	GTGTTTCTTG	CCCAGGAAAT	ACTACGACTG	2460
	ACTTTGATGG	CTCCACAAAC	ATAACCCAGT	GTAAAAACAG	AAGATGTGGA	GGGAGGCTGG	2520

5
10
15
20

```

GAGATTTTAC TGGGTACATT GAATCCCCAA ACTACCCAGG CAATTACCCA GCCAACACCG 2580
AGTGATACGTG GACCATCAAC CCACCCCCCA AGCGCCGCAT CCTGATCGTG GTCCCTGAGA 2640
TCTTCTCTGCC CATAGAGGAC GACTGTGGGG ACTATCTGGT GATGCGGAAA ACCTCTTCAT 2700
CCAAATTCGTG GACACATATAT GAAACCTTGGC AGACCTACGA ACGCCCATC GCCTTCACCT 2760
CCAGGTCAAA GAAGCTGTGG ATTCAGTTCA AGTCCAATGA AGGGAACAGC GCTAGAGGCT 2820
TCCAGGTCCC ATACGTGACA TATGATGAGG ACTACCAAGG ACTCATGAA GACATAGTTC 2880
GAGATGGCAG GCTCTATGCA TCTGAGAACC ATCAGGAAAT ACTTAAGGAT AAGAAACTTA 2940
TCAAGGCTCT GTTGTATGTC CTGGCCCATC CCCAGAACTA TTTCAAGTAC ACAGCCGAGG 3000
AGTCCCGAGA GATGTTTCCA AGATCGTTCA TCCGATTGCT ACCTTCCAAA GTGTCCAGGT 3060
TTTTGAGACC TTACAAATGA CTCAGGCCAC GTGCCACTCA ATACAAATGT TCTGCTATAG 3120
GGTTGGTGGG ACAGAGCTGT CTCTCTCTCG CATGTGAGCA CAGTCGGGTA TTGCTGCCTC 3180
CGGTATCAGT GACTCATTAG AGTTCAATTT TTATAGATAA TACAGATATT TTGTTAAATT 3240
GAACITGGTT TTTCTTTCCC AGCATCGTGG ATGTAGACTG AGAATGGCTT TGAGTGGCAT 3300
CAGCTTCTCA CTGCTGTGGG CGGATGTCCT GGATAGATCA CGGCTGGCT GAGCTGGACT 3360
TTGGTCAGCC TAGGTGAGAC TCACCTGTCC TTCTGGGGTC TTACTCTCTC TCAAGGAGTC 3420
TGTAGTGAA AGAGGGCCAC AGAATAAGCT GCTTATCTCG AAACCTTCAGC TTCTCTAGC 3480
CCGCCCCCTC CTAAGGGAGC CTTCTGCAC CTCTGACAG CTTGACCCAG GCAGAACAGG 3540
CAAGAGGGGA GGGAAAGAGA CCCTGCAGG CTCCCTCCAC CCACCTGAG ACCTGGGAGG 3600
ACTCAGTTTC TCCACAGCCT TCTCCAGCCT GTGTGATACA AGTTTGATCC CAGGAACCTG 3660
AGTTCTAAGC AGTCTCGTG AAAAAAANA GCAGAAAGAA TTAGAAATAA ATAAAACTA 3720
AGCATTCTG GAGACAT

```

Seq ID NO: 52 Protein sequence

Protein Accession #: NP_066025

25
30
35
40
45

```

1 11 21 31 41 51
MGVAGNRNPG AAWAVLLLLL LLPLLLLLAG AVPPGRGRRA GPQEDVDECA QGLDDCHADA 60
LCQNTPTSYK CSCKPGYQGE GRQCEDIDEC GNELNGGCVH DCLNIPGNRY CTCPDGFMLA 120
EDGHNLCDVD ECLNNNGGCG HTCVNVMGSS ECCCCGFFL SDNQHTCYHR SEBGLSCMNK 180
DHGCSHICKE APRGSVACBC RPFGEELAKNQ RDCILTCNHG NGGCOHSCDD TADGPBCSCH 240
PQYKMTDGR SCLEREDTVL EVTESNTT9V VDGDKRVKRR LLMETCAVNN GGDRTCKDT 300
STGVHCSCPV GFTTLQDQKT CKDIDECQTR NGGCDHFCKN IVGSPDCGCK KGFKLITDEK 360
SCQDVDECSL DRTCDHSCIN HPGTFAACAN RGYTLYGFTH CGDTNECSIN NGGCGQVCVN 420
TVGSYECQCH PGYKLHWNKK DCEVEKGLLP TSVSPRVSLH CGKSGGGDGC FLRCHSGIHL 480
SSDVTITRTS VTFFCLNKGCK SLKNALPPE GLRPALEPKE SSVKESFRIV NLTCSSGKQV 540
FGAGRPSTP KEMFTITVFE LETNQKEVTA SCDLSCIVKR TEKRLRKAIR TLRKAVHREQ 600
FHLQLSGMNL DVAKKPPRTS BRQAESCGVG QGHAENQCVS CRAGTYDGA RERCILCPNG 660
TFQNEGGQMT CSECPFRPNS GALKTPEAWN MSECGLCOP GREYADGFAP CQLCALGTPQ 720
PERGRTSCFP CGGGLATKQ GATSFQDCET EVQCSFGEFY NTTTHRCIRC PVGTYQPEFG 780
KNKCVSCFEN ITTDEGDSIN ITQCKNRKCG GELGDFTYI ESPNYPGNYF ANTECTWTIN 840
PFPKRLILV VPEIFLPIED DGDYLVMRK TSSNSVTTY ETCQTYERPI APTSRSEKLV 900
IQPKSNEGNS ARGFQVPYVT YDEYQELIE DIVRDGLYA SENHQEILKD KKLKALFDV 960
LAHPQNYFKY TAQBSREMF RSPIRLLRSK VSRFLRPYK

```

Seq ID NO: 53 DNA sequence

Nucleic Acid Accession #: NM_014211

Coding sequence: 157..1479

50
55
60
65
70
75
80

```

1 11 21 31 41 51
GGGACAGGGC TGAGGATGAG GAGAACCTCG GGGACCCAGA AGACCGTGCC TTGCCCGGAA 60
GTCCTGCCTG TAGGCTTGAA GACCTTGCCC TAACAGAGCC TCAACAACCTA CCTGGTGATT 120
CCTACTTCTG CCGCTTGGTG TGAGCAGCTT CTCAACATGA ACTACAGCCT CCCTCTGGCC 180
TTCTGTGTCT TGAGTCTCTT CACTGAGAGG ATGTGCATCC AGGGGAGTCA GTTCAACGTC 240
GAGGTGCGCA GAAGTAGCAA GCTTTCCTCG CTTGGCTTTG AGAACCTCAC AGCAGGATAT 300
AACAAATTTT TCAGGCCCAA TTTTGGTGGA GAACCGTAC AGATAGCGCT GACTCTGGAC 360
ATTGCAAGTA TCTCTAGCAT TTCAGAGAGT AACATGGACT ACACAGCCAC CATATACCTC 420
CGACAGCGCT GGAATGACCA GCGGCTAGTG TTTGAAGGCA ACAGAGCTT CACTCTGGAT 480
GCCCGCCTCG TGGAGTTCTT CTGGGTGCCA GATACTTACA TTGTGGAGTC CAAGAAGTCC 540
TTCTCTCCTG AAGTCACTGT GGGAAACAGG CTCATCCGCC TCTTCTCCAA TGGCACGGTC 600
CTGTATGCCC TCAGAACTAC GACAACTGTT GCATGTAACA TGGATCTGTC TAAATACCCC 660
ATGACACACAC AGACATGCAA GTTGACCTG GAAGCTGGG GCTATGATGG AAAATGATGT 720
GAGTTCACCT GGCTGAGAGG GAACGACTCT GTGCGTGGAC TGGAACACCT GCGGCTTGCT 780
CAGTACACCA TAGAGCGGTA TTTCACTTCA GTCAACAGAT CGCAGCAGGA GACAGGAAAT 840
TACACTAGAT TGGTCTTACA GTTTGAGCTT CGGAGGAATG TTCTGTATTT CATTTTGGAA 900
ACCTACGTTT CTTCACCTTT CTTGGTGGTG TTGTCTGGG TTTTATTITG GATCTCTCTC 960
GATTCACTCC CTGCAAGAAC CTGCATTGGA GTGACGACCG TGTATCAAT GACCACACTG 1020
ATGATCGGGT CCGCACCTTC TCTTCCCAAC ACCAAGTGT TCATCAAGGC CATCGATGTG 1080
TACCTGGGGA TCTGCTTTAG CTTTGTGTTT GGGGCTTGG TAGAATATGC AGTTGCTCAC 1140
TACAGTTCTT TCAGCAGAT GGCAGCCAAA GATAGGGGGA CAACAAGAA AGTAGAAGAA 1200
GTCAATATTA CTAATATCAT CAACAGCTCC ATCTCCAGCT TTAACCGGAA GATCAGCTTT 1260
GCCAGCATTG AAATTTCCAG CGACACGTT GACTACAGTG ACTTGACAAAT GAAAACCCAGC 1320
GACAGTTTCA AGTTTGTCTT CCGAGAAAAG ATGGGCAGGA TTGTTGATTA TTTCACAATT 1380
CAAAACCCCA CTAATGTTGA TCACTATTCC AAACACTGT TCTCTTTGAT TTTTATGCTA 1440
GCCAATGTAT TTTACGTGGC ATACTACATG TATTTTGTAG TCAATGTTAA ATTTCTTGCA 1500
TGCCATAGGT CTTCAACAGG ACAAGATAAT GATGTAAATG GTATTTTAGG CCAAGTGTGC 1560
ACCCACATCC AATGGTGCTA CAAGTACTG AAATAATATT TGAGTCTTTC TGCTCAAGAA 1620
ATGAGCTCC AACCATTTT CTAAGCTGTG TAGAAGTCTT AGCATATAG GATCTGTATA 1680
TAGAAACATC AGTCCATTCC TCTTTCATCT TAATCAAGGA CATTCOCATG GAGGCCAAGA 1740
TTACAAATGT ACTCAGGGCT GTTTATTGCG TGGCTCCCTG GTTGTGATTT ACCTCATATA 1800
AAGAAATGGA AGGAGACCAT TGGGTAAACC TCAAGTGTCA GAAGTTGTTT CTAAGTAAAC 1860
TATACATGTT TTTTACTAAA TCTCTGAGT GCTTATATAA TACATTGTTC CTAATTAGG 1920
GAGTAACATT TTCTAGTTTT TGTTCCTGGT TAAATGAAA TATGGGCTTA TGTCAATTCA 1980

```


5 TTGGAACTCA ATGCACTAAC TCAATACCAA GATGAGTTT TAAATAATGA ATATTATTTA 2040
 ATACCAACAC AGAATTATCC CCAATTTCCA ATAAGTCTTA TCATTGAAAA TTCAATATATA 2100
 AGTGAAGAAA AATTTAGTAG ATCAACAATC TAAACAAATC CCTCGGTTCT AAGATACAAT 2160
 GGATTCCOCCA TACTGGAAGG ACTCTGAGGC TTATTTCCCC CACTATGCAT ATCTTATCAT 2220
 TTTATTATTA TACACACATC CATCCTAAAC TATACTAAAG CCTTTTCC ATGCGATGGAT 2280
 GGAATGGAA GATTTTFTTG TAATTTGTTT TAGAAGTCTT AATATGGGCT GTTGCCATGA 2340
 AGGCTTGCGG AATTGAGTCC ATTTCTTAGC TGCCTTTATT CACATAGTGA TGGGGTACTA 2400
 AAAGTACTGG GTTGACTCAG AGAGTGGCTG TCATTCGTG ATTGCTGCTA CTCTAACACT 2460
 GAGCAACACT CTCCAGTGG CAGATCCCTT GTATCATTC AAGAGGAGCA TTCATCCCTT 2520
 TGCTCTAATG ATCAGGAATG ATGCTTATTA GAAAACAAAC TGCTTGACCC AGGAACAAGT 2580
 GGCTTAGCTT AAGTAAACTT GGCTTTGCTC AGATCCCTGA TCCTTCCAGC TGGTCTGCTC 2640
 TGAGTGGCTT ATCCCGCATG AGCAGGAGCG TGCTGGCCCT GAGTACTGAA CTTCCTGAGT 2700
 AACAAATGAG CAGTTTACAG AACCTATGTT CAGGTGCGG GTGAGCTGCC CTCTCCAAAT 2760
 CCAGCCAGAG ATGCACATTC CTGCGCCAGT CTCAGCCAAAC AGTACCAAAA GTGATTTTGT 2820
 AGTGTGGCCG GGTAAAGGCT TCAGTTCAG CCTCAGTTAT TTTAGACAAAT CTCGCCATCT 2880
 TTAATTTCTT AGCTTCTGT TCTAATAAAT GCACGGCTTT ACCTTTCTCT TCAGAAATAA 2940
 ACCAAGGCTC TAAAGATGA TTTCCTTCT GTAACTCCCT AGAGCCACAG GTTCTCATTC 3000
 CTTTTCCCAT TATATCTCTC ACAATTCAGT TTCTATGAGT TTGATCACCT GATTTTITTA 3060
 ACAGAAATAT TCTAACGGGA ATGGGTGGGA GTGCTGGTGA AAGAGATGA AATGTGGTGT 3120
 20 TCTAGGCCAA TCTATTTTGT GATTTTITTA AAAAAGTTTA AAGGAATAA TCTGTCTCTA 3180
 AACCCACTT AAGCATTTGT TTTATATAAA AACAAATGATA AAGATGTGAA CTGTGAAATA 3240
 AATATACCAT ATTAGCTACC CACC

25 Seq ID NO: 54 Protein sequence
 Protein Accession #: NP_055026

30 1 11 21 31 41 51
 MNYSLHLAEV CLSLPETERMC IQGSGPNVEV GRSDKLSLPG FENLTAGYNK FLRPNFGGEP 60
 VQIALTLIDIA SISISSESNM DYTATYLRQ RMDQLRVE GNKSFILDR LVEFLWVPT 120
 YIVEKKKSL HEVTVGNRI RLPNGTIVLY ALRITTTVAC NMDLSKYEMD TQCKLQLES 180
 WGYDNDVEF TWLRGNDVSR GLEHLRLAQY TIERVFTLVY RSQGETGNYT RLVIQPELR 240
 NVLYFILSY VPSTFLVVL VWSFWISLDS VPARTICIGVT TVLSMTILMI GSRTSLPNT 300
 CFIKALDVYL GICPSFVFGA LLEYAVAYES SLQQMAAKDR GTTKEVEEVS ITWINSSTIS 360
 35 SFKRKLSPAS LEISSDNVDY SDLTMTSDK FKEVFEKMG RIVDYFTIQN PSNVDRHYSKL 420
 LPLLIPLMLAN VFYWAYMYF

40 Seq ID NO: 55 DNA sequence
 Nucleic Acid Accession #: XM_084007
 Coding sequence: 138..2405

45 1 11 21 31 41 51
 CTGCTGCCGA ATTCGGCAGC AGACCGCGTG TTCGCGCCTG GTAGAGATT TCGAAGACA 60
 CCACTGGGCC CGPTGGGAAC CAAACCTGGG OCGGTGGCGG GCGCGTGGGA CACGAGGCC 120
 GCGGAGACGA AGGCGCAATG GCGAGGAAGT TATCTGTAAT CTGTATCTG ACCTTTGGCC 180
 TCTCTGTCTC AATCCCTCTT CATGAACATA AAGCAGCTGC TTTCCGCCAG ACCACTGAGA 240
 AAATTAGTCC GAATTGGGAA TCTGGCATTG ATGTTGACTT GGCAATTTCC ACAGGGCAAT 300
 50 ATCATCTACA ACAGCTTTTC TACCGCTATG GAGAAATAA TTCTTTGTC GTTGAAGGGT 360
 TCAGAAAAAT ACTTCAAAAT ATAGGCTATG ATAAGATTAA AAGAATCCAT ATACACCATG 420
 ACCACGACCA TCATCTCAGC CACGAGCATC ACTCAGACCA TGAGCTCTAC TCAGACCATG 480
 AGCATCACTC AGACCAAGAG CATCACTCTG ACCATGATCA TCACCTCCAC CATAATCATC 540
 CTGCTTCTGG TAAAAAATAG CGAAAAGCTC TTTGCCGAGA CCATGACTCA GATAGTTCTG 600
 55 GTAAGATCC TAGAACAGC CAGGGGAAAG GAGCTCACCG ACCAGAACAT GCCAGTGGTA 660
 GAAGGAATGT CAGGACAGT GTTAGTCTTA GTGAAGTGC CTCAACTGTG TACAACACTG 720
 TCTCTGAAGG AACTCACTTT CTAGAGACAA TAGAGACTCC AAGACCTGGA AAACCTCTCC 780
 CCAAGATGT AAGCACTCC ACTCCACCCA GTGTCAATC AAAGAGCCGG GTGAGCCGGC 840
 TGGCTGCTAG GAATCAAAAT GAATCTGTGA GTGAGCCCGG AAAAGGCTTT ATGTATTCCA 900
 60 GAAACACAAA TGAAAATCCT CAGGAGTGT TCAATGCATC AAAGCTACTG ACATCTCATG 960
 GCATGGGCAT CAGGTTCCG CTGAATGCAA CAGAGTTCAA CTATCTCTGT CCAGCCATCA 1020
 TCAACCAAAAT TGATGCTAGA TCTGTCTGTA TTCATACAAG TGAAAAGAG GCTGAAATCC 1080
 CTCCAAAGAC CTATTCATTA CAAATAGCCT GGGTGGTGG TTTTATAGCC ATTTCCATCA 1140
 TCAGTTTCTT GTCTCTGCTG GGGGTATCT TAGTGCTCT CATGAATCGG GTGTTTTTCA 1200
 65 AATTTCTCCT GAGTTTCTCT GTGGCACTGG CCGTTGGGAC TTTGAGTGGT GATGCTTTTT 1260
 TACACCTTCT TCCACATTCT CATGCAAGTC ACCACCATAG TCATAGCCAT GAAGAACCAG 1320
 CAATGGAATG GAAGAAGGGA CCACCTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAAA 1380
 GTGCCATTIT TGATTCACG TGAAGGGTTC TAACGCTCT AGGAGGCTG TATTTCTGTT 1440
 TTCTTGTGTA ACRTGTCTTC ACATTGATCA AACATTTTAA AGTAAGAAG AAAAAGAATC 1500
 70 AGRAGAAACC TGAATATGAT GATGATGTGG AGATTAAAG GCAGTTGTCC AAGTATGAAT 1560
 CTCACACTTC AACAAATGAG GAGAAAGTAG ATACAGATGA TCGAACTGAA GGCTATTATC 1620
 GAGCAGACTC ACAGAGAGCC TCCCACTTTG ATCTCTCAGCA GCCTGCAGTC TTGGAAGAG 1680
 AAGAGTCTAT GATAGCTCAT GCTCATCCAC AGGAAGTCTA CAATGAATAT GTACCCAGAG 1740
 GGTGCAAGAA TAAATGCCAT TCACATTTCC ACGATACACT OGGCCAGTCA GACGATCTCA 1800
 75 TTCACACCCA TCATGACTAC CATCATATTC TCCATCATCA CCACCAACCA AACCAACTAC 1860
 CTCACACTCT CAGCCAGGCG TACTCTCGGG AGGAGCTGAA AGATGCGGGC GTGCCACTT 1920
 TGGCTCTGAT GGTGATAATG GGTGATGGCC TGCACAAATT CAGCBAATGC CTAGCAATTT 1980
 GTGCTGCTTT TACTGAAGGC TTATCAAGTG GTTTAAGTAC TTCTGTGTGT GTGTTCTGTC 2040
 ATGAGTTGCC TCATGAATTA GGTGACTTTG CTGTTCTACT AAGGCTGTC ATGACCGTTA 2100
 80 AGCAGGCTGT CTTTATAAT GCATTGTCTG CCATGCTGTC GTATCTTGGG ATGGCAACAG 2160
 GAATTTTCTT TGTCTATTAT TTTCTATGTG GATATTGCA CTACCTGCTG 2220
 GCTTATTATC GTATGTTGCT CTGGTTGATA TGGTACCTGA AATGCTGCAC AATGATGCTA 2280
 GTGACCATGG ATGTAGCCGC TGGGGGTATT TCTTTTACA GAATGCTGGG ATGCTTTTGG 2340
 GTTTTGGAAAT TATGTTACTT ATTTCCATAT TTGAACATAA AATGCTGTTT CGTATAAATT 2400
 TCTAGTTAAG GTTTAAATGC TAGAGTAGCT TAAAGAGTTG TCATAGTTTC AGTAGTTCAT 2460

5
10
15

```

AGGGAGATGA GTTGTATGTC TGTACTATGC AGCGTTTAAA GTTAACTGGGT TTTGTGATTT 2520
TTGTATTGAA TATTGCTGTC TGTACAAAG TCAGTTAAAG GTACGTTTAA ATATTTAAGT 2580
TATTCTATCT TGGAGATAAA ATCTGTATGT GCAATTCACC GGTATTACCA GTTTATTATG 2640
TAAACAAGAG ATTTGGCATG ACATGTTCCTG TATGTTTCAG GGAATAATGT CTTTAAATGCT 2700
TTTTCAAGAA CTAAACACAGT TATTCCTATA CTGGATTTTA GGTCTCTGAA GAACGTCTGG 2760
TGTTTAGGAA TAAGAATGTG CATGAAGCCT AAAATACCAA GAAAGCTTAT ACTGAATTTA 2820
AGCAAAGAAA TAAGGAGAAA AAGAGAAAG TCTGAGAAAT GGGAGGCCAT AGATTCTTAT 2880
AAAAATCACA AAATTTGTTG TAAATTAGAG GGGAGAAATT TAGAATTAAG TATAAAAAGG 2940
CAGAATTAGT ATAGAGTACA TTCATTAAAC ATTTTGTGCA GGATTATTTC CCGTAAAAC 3000
GTAGTGAGCA CTCTCATATA CTAATTAGTG TACATTAAAC TTTGTATAAT ACAGAAATCT 3060
AAATATATTT AATGAATTCA AGCAATATAC ACTTGACCAA GAAATTGGAA TTTCAAATG 3120
TTCGTGCGGG TTATATACCA GATGAGTACA GTGAGTAGTT TATGTATCAC CAGACTGGGT 3180
TATTGCCAAG TTATATATCA CCAAAGCTG TATGACTGGA TGTCTGGTT ACCTGGTTA 3240
CAAAATTATC AGAGTAGTAA AACTTTGATA TATATGAGGA TATTAAACT ACACAAAGTA 3300
TCATTTGATT CCAATCAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCATTTG 3360
GAGCAATGTG CTTTATATAC GGTACTGTAG CCATACTAGG CCGTCTGTG GCATTCCTTA 3420
GATGTTTCTT TTTTACAAA TAATTCCTT ATATCAGCTT G

```

Seq ID NO: 56 Protein sequence

Protein Accession #: XP_084007

20
25
30
35

```

1 11 21 31 41 51
| | | | |
MARKLSVILI LTFALSVTNP LHBLKAAAF QTTEKISPNN ESGINVDLAI STROYHLQQL 60
PYRYGENNLS SVEGFRLLQ NIGIDKIKRI HIHHDMDHHS DHEHSDHER HSDHREHSDH 120
BHHSDHHS HHNHAASGKN KRIALCFDHD SDSSGKDPNN SQKGGAHRPE HASGRNVD 180
SVSASEVIST VYNTVSEGTN FLETIETPRP GKLFPRIVSS STPPSVTSKS RVSRLAGRKT 240
NESVSEPRRG FMYSRNTNEN PQECFNASKL LTSHGMGIQV PLNATEFNYL CPAINQIDA 300
RSLCLHTSRK KAEIPPKTVS LQIAWVGFI AISIISFLSL LGVILVPLMN RVFFKFLSP 360
LVLAVALTSL GDALFHLLEH SHASHHSHS HEEPAMEMKR GPLFSLSSQ NIEESAYFDS 420
TWKGLTALGG LYFMFLVEHV LTLIKQFKDK KKKNQKKPEN DDDVEIKQL SKYESQLSTN 480
EEKVDTDRT EGYLRADSQE PSHFDSQQA VLZEEVMIA HAHPEVYNE YVPRGCKNKC 540
HSHFHDITLQ SDDLIEHHHD YHILHSHH QNHHPHSHSQ KYSREELKDA GVATLAWMI 600
MDEGLHNPDS GLATGAATFE GLSSGLSTSV AVFCHLPHE LGDFAVLKA GMTVKQAVLY 660
NALSAMLAYL GMATGIFIGH YAEVSMWIF ALTAGLFMYV ALVDMVPEML HNDASDHGCS 720
RWGYFPLQNA GMLGFGIML LISIFHKIV FRINF

```

Seq ID NO: 57 DNA sequence

Nucleic Acid Accession #: NM_015419.1

Coding sequence: 1..8487

40
45
50
55
60
65
70
75
80

```

1 11 21 31 41 51
| | | | |
ATGCCCAAGC GCGCCCACTG GGGGGCCCTC TCGTGGTGC TGATCCTGCT TTGGGGCCAT 60
CCGCGAGTGG CCGTGGCCTG CCGGCATCCT TGTGCCTGCT ACCTCCCAAG CGAGGTCCAC 120
TGCACATTCG GATCCCTGCG TCCGTGCCCC GCTGGCATTG CTAGACACGT GGAAGAATC 180
AATTTGGGGT TTAATAGCAT ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG 240
TTGAGCTAC TTATGATTC ACGCAATGAG ATCCCAAGCA TCCCGATGG AGCTTTAAGA 300
GACCTCAGCT CTCCTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT GATCACAGGA 360
CAGACCTCC AGGGCTCTC TAACTTAATG AGGCTGCACA TTGAACACAA CAAGATCGAG 420
TTTATCCACC CTCAGCTTT CAACGGCTTA AGCTCTCTGA GGCTACTCCA TTTGGAAGGA 480
AACTCCTCC ACCAGCTGCA CCCCAGCACC TTCTCCAGT TCACATTTT GATTATTTC 540
AGACTCTCCA CCAATAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCTGCC 600
AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAT TCCGTGGACC 660
TGGCATTTG TGAGATGATG GTTTTGGAA TGGGATGCAA AATCCAGAG AATCTGAGG 720
TGTAAGAAAG ACAAGCTTA TGAAGGCGGT CAGTTGTGTG CAATGTGCTT CAGTCCAAAG 780
AAGTTGTACA AACATGAGT ACACAGCTG AAGGACATGA CTGTCTGAA GCCITCAATA 840
GAGTCCCTTC TAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG 900
GATGTGGCA GCCAGCTCAT CCGTGAGAAA TTCCAATGCG CCCAGTGGAG CATCTCTTG 960
AATATGACCG ACAGACAGCG GAACATGCTG AACTTGGTCT GTGACATCAA GAAACCAATG 1020
GATGTGTACA AGATTCACTT GAAACAAAG GATCCTCCAG ATATTGACAT AATGCAACA 1080
GTTCCTTGG ACTTTGAGTG TCCATGACC CGAGAAACCT ATGAAAGCT ATGGAAATG 1140
ATAGCATACT ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAGAC 1200
CCCAGAGTCA GCTACAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA CTACACAGGT 1260
GTGAGAGCCC AGATTCTTGC AGAACCAAG TGGGTCAATG AGCCATCCAT AGATATCCAG 1320
CTGACCCGAC GTACAGATAC GGCCAAGAG GTGCTACTTT CCTACTACAC CCAGTATCT 1380
CAACAATAT CCACCAAGA TACAAGCCAG GCTCGGGA CAAGCTGGGT AATGATTGAG 1440
CCTAGTGGAG CTGTGCAAG AGATCAGACT GTCTGGAAG GGGTCCATG CCAGTTGAGC 1500
TGCAACGTGA AACCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCAGA TGGCTCCATC 1560
CTGAAGCCTC CCAATGATGA CCCAGACAGC AAGTTCTCCA TTCTCAGCAG TGGCTGGCTG 1620
AGGATCAAGT CCAATGAGCC ATCTGACTCA GGCTGTGACC AGTGCATTGC TCAAGTGAAG 1680
GATGAATGG ACCGATGGT ATATAGGTA CTGTGCAAT CTCCCTCCAG TCAGCCAGCC 1740
GAGAAAGACA CAGTGACAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT 1800
GCTTTAGCAA TACCCGAGC CCACTTAGC TGGATTCTTC CAACAGAGG GATAATTAA 1860
GATTGTGCTA ACACATCACA TGTATACATG TTGCCAATG GAACTCTTTC CATCCCAAG 1920
GTCCAGTCA GTGATAGTGG TTACTACAGA TGTGTGGCTG TCAACAGCA AGGGGCAGAC 1980
CATTTTACGG TGGGAATCAC AGTGACCAAG AAAGGCTCTG GCTTGCTATC CAAGAGGC 2040
AGACGCCAG GTGCAAGGC TCTTTCAGA GTACAGAGAG ACATCGTGA GGATGAAGG 2100
GGCTGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC TTCTGCATCC AAGGACCAA 2160
GAGGTGTTCC TCAAAACAA GGATGATGCC ATCAATGGAG ACAAGAAAGC CAAGAAAGG 2220
AGAGAAAGC TGAACCTCTG GAAGCATTCG GAAAGAAAGC CAGAGACCA TGTTCAGAA 2280
GGTCCAGAG TGTTTGAATC TAGACGAGG ATAAACATGG CAACAAACA GATTAACTCG 2340
GAGCGCTGG CTGATATTTT AGCCAAAGTC CGTGGGAAA ATCTCCCTAA GGGCAGAA 2400
GTACCCCGT TGATTAAAC CACAAGTCTT CCATCCTTGA GCCTAGAAAT CACACCACCT 2460

```

	TTCCTGCTG	TTTCTCCCC	CTCAGCATCT	CCTGTGCAGA	CAGTAAOCAG	TGCTGAAGAA	2520
	TCCTCAGCAG	ATGTACTCT	ACTTGGTGAA	GAAGAGCACG	TTTGGGTAC	CATTTCCTCA	2580
	GCCAGCATGG	GGCTAGAAC	CAACCAAA	GGAGTTATTC	TTGTTGAACC	TGAAGTAACA	2640
5	AGCAGCATCT	TGGAGGAAGT	TGTTGATGAC	CTTTCGAG	AGACTGAGGA	GATAACTTCC	2700
	ACTGAAGGAG	ACCTGAAGGT	GACAGCAGCC	CCTACACTTA	TATCTGAGCC	TTATGAACCA	2760
	TCTCTACTIC	TGCACACATT	AGACACAGTC	TATGAAAAGC	CCACCCATGA	AGAGACGGCA	2820
	ACAGAGGGTT	GGTCTGCAGC	AGATGTTGGA	TCGTCAACAG	AGCCCAATC	CAGTGAGTAT	2880
	GAGCCTCCAT	TGGATGCTGT	CTCCTTGGCT	GAGTCTGAGC	CCATGCAATA	CTTTGACCCA	2940
10	GATTTGGGAG	CTAAGTCACA	ACCAGATGAG	GATAAGATGA	AAGAAGACAC	CTTTGCACAC	3000
	CTTACTCCAA	CCCCACCAT	CTGGGTAAAT	GACTCCAGTA	CATCACAGTT	ATTGAGGAT	3060
	TCTACTATAG	GGGAACAGG	TGTCACAGCC	CAATCACATC	TACAAGGACT	GACAGACAAC	3120
	ATCCACCTTG	TGAAAAGTAG	TCTAAGCACT	CAAGACACCT	TACTGATTAA	AAAGGGTATG	3180
	AAAGAGATGT	CTCAGACACT	ACAGGGAGGA	AATATGCTAG	AGGGAGACCC	CACACACTCC	3240
15	AGAAGTCTTG	AGAGTGAGGG	CCAGAGAGGC	AAATCCATCA	CTTTGCCCTGA	CTCCACACTG	3300
	GGTAATAAG	AGAGTGAAG	TCCAGTTAAG	AAGCCTGCGG	AAACCAAGT	TGTTACCCCTC	3360
	CTAGACAAAG	ACACCAACAAC	AGTAACAACA	ACAACCAAGC	AAAAAGTTGC	TCCGTCTATC	3420
	ACCATGAGCA	CTCACCCCTC	TGGAAGGAGA	CCCAACGGGA	GAAGGAGATT	ACGCCCCAAAC	3480
	AAATTCGCC	ACCGGCACAA	GCAAAACCCA	CCCACAACCT	TTGCCCCATC	AGAGACTTTT	3540
20	TCTACTCAAC	CAACTCAAGC	ACCTGACATT	AAGATTTCAA	GTCAAGTGG	GAGTTCTCTG	3600
	GTTCCTACAG	CTTGGGTGGA	TAAACAGTT	AATACCCCA	AACAGTTGGA	AATGGAGAAG	3660
	AATGAGAAC	CCACATCCAA	GGGAACACCA	CGGAGAAAC	ACGGGAAGAG	GCCAAACAAA	3720
	CATCGATATA	CCCTTTCTAC	AGTGAGCTCA	AGAGCGTCCG	GATCCAGCC	CAGCCCTTCT	3780
	CCAGAAATA	AACATAGAAA	CATTGTTACT	CCCAGTTGAG	AAACTACTACT	TTTGGCTAGA	3840
25	ACGTGTTCTC	TGAAAACCTGA	GGGCCCTTAT	GATTCCTTAG	ATTACATGAC	AACCAACAGA	3900
	AAAAATATAT	CATCTATCCC	TAAAGTCCAA	GAGACACTTC	CAGTCACATA	TAAACCCACA	3960
	TCAGATGGAA	AAGAAATTA	GGATGATGTT	GCCACAATG	TTGACAAACA	TAAAGTGAC	4020
	ATTTAGTCA	CTGGTGAATC	AATTACTAAT	GCCATACCAA	CTTCTCGCTC	CTTGGTCTCC	4080
	ACTATGGGAG	AATTTAAGGA	AGAATCCTCT	CCTGTAGGCT	TTCCAGGAAC	TCCAACTCGG	4140
30	AATCCCTCAA	GGAGCGGCCA	GCCTGGGAGG	CTACAGACAG	ACATACCTGT	TACCACTTCT	4200
	GGGAAATC	TTACAGACCC	TCCCTTCTT	AAAGAGCTTG	AGGATGAGGA	TTTCACTTCC	4260
	GAGTTTGT	CCTCTTGAC	AGTCTCCACA	CCATTTCACC	AGGAGAGAGC	TGGTCTTCC	4320
	ACAACCTCTC	CAAGCATAAA	AGTGGAGGTG	GCCTCAAGTC	AGGCAGAAC	CACCACCTTT	4380
	GATCAAGATC	ATCTGAAAC	CACCTGAGCT	ATTCTCCTTT	CTGAAACTAG	ACCACAGAT	4440
35	CACACCCCTA	CTGCTGCGCG	GATGAGGAG	CCAGCATCCT	CGTCCCATC	CACAATCTCT	4500
	ATGTCTTGG	GACAAACAC	CACCACTAAG	CCAGCACTTC	CCAGTCCAG	AATATCTCAA	4560
	GCATCTAGAG	ATTCAGAGGA	AAATGTTTTT	TJGAATATG	TGGGGAATCC	AGAAACAGAA	4620
	GCACCCCTAG	TCACCAATGA	AGGAACACAG	CATATGTCAG	GGCCAAATGA	ATTATCAACA	4680
	CCCTCTTCOG	ACCGGAGTGC	ATTAACTTG	TCTACAAAGC	TGGAATTGGA	AAAGCAAGTA	4740
40	TTTGGTAGTA	GGAGTCTACC	ACGTGGGCCA	GATAGCCAAC	GCCAGGATGG	AAGAGTTTAT	4800
	GCCTCTCATC	AACAAACAG	AGTCCCTGCC	AAACCCATCC	TACCAACASC	AACAGTGAGG	4860
	CTACCTGAAA	TGTCCACACA	AAGCGCTTCC	AGATACCTTG	TAACTTCCCA	GTCACTCTGT	4920
	CACCTGGACA	ACAAACCGGA	AATAACTACA	TATCCTTCTG	GGCCTTTGCC	AGAGAACAAA	4980
	CAGTTTACAA	CTCCAGATT	ATCAAGTACA	ACAAITCCTC	TCCCATTTGA	CATGTCCAAA	5040
45	CCAGCATTC	CTAGTAAGTT	TACTGACCGA	AGAACTGACC	AATTCATGCG	TTACTCCAAA	5100
	GTGTTTGGAA	ATAACCAATC	CCCTGAGGCA	AGAAACCCAG	TGGGAAAGCC	TCCCACTCCA	5160
	NGAATTCCTC	ATTATTCCAA	TGGAGAGCTC	CCTTCTTTA	CCAACAAGAC	TCTTCTTTT	5220
	CCACAGTTGG	GAGTCAACCG	GAGACCCCG	ATACCCACTT	CTCTGCCCC	AGTAATGAGA	5280
	GAGAGAAAAG	TTATTCAGG	TTCTTACAAC	AGGATACATT	CCCATAGCAC	CTTCAATCTG	5340
50	GACTTTGGCC	CTCCGACACC	TCCGTTGTTG	CACACTCCGC	AGACCCACGG	ATCACCTTCA	5400
	ACTAATCTAC	AGAACTATCC	TATGGTCTCT	TCCACCCAGA	GTTCATCTCT	CTTTATAACA	5460
	TCTTCTGTCC	AGTCTCTGCG	AAGCTTCCAC	CAGAGCAGCT	CAAAGTTCTT	TGCAGGAGGA	5520
	CCTCCGCTAT	CCAAATTTCT	GTCTCTTGGG	GAAAAGCCCC	AAATCCTCAC	CAAGTCCCCA	5580
	CAGACTGTGT	CCGTCAACCG	TGAGACAGAC	ACTGTGTTC	CCTGTGAGGC	AACAGGAAAA	5640
55	CCAAAGCCCT	TGTTTACTTG	GACAAAGGTT	TCCACAGGAG	CTCTTATGAC	TCCGAATACC	5700
	AGGATACATC	GGTTTGAAGT	TCTCAAGAAC	GSTACCTTAG	TGATACGGAA	GGTTCAAGTA	5760
	CAAGATCAGG	GCCAGTATAT	GTGCACCCCT	AGCAACTGCG	ACGGCTCTGA	CAGGATGGTG	5820
	GTCTGTCTTT	CGGTCAACCG	GCAACACCTT	CAATCTCTAG	CCTCCCACTA	CCAGGACGTC	5880
	ACTGTCTACC	TGGGAGACAC	CATTGCAATG	GAGTGTCTGG	CCAAAGGGAC	CCAGCCCCC	5940
60	CAAAITTCCT	GGATCTTCCC	TGACAGGAGG	GTGTGGCACA	CTGTGTCCCC	CGTGGAGAGC	6000
	CGCATCACCC	TGCACGAAAA	CCGACCCCTT	TCATCAAGG	AGGCGTCTTT	CTCAGACAGA	6060
	GGCGTCTATA	AGTGGGTGGC	CAGCAATGCA	GCGGGGGCGG	ACAGCCTGGC	CATCCGCTCG	6120
	CAGGTGGCGG	CAGTGGCCCC	CGTTATCCAC	CAGGAGAACG	TGGAGAACAT	CTCGCTGCCG	6180
	CCGGGGCTCA	GCATTCACAT	TCAGTGCATC	GCCAGGCTG	CGCCCCCTGC	CAGGCTGCCG	6240
65	TGGGTGCTCG	GGGAGGGTAC	CCAGATCCGC	CCCTGCGAGT	TCCTCCACGG	GAACTTGTIT	6300
	GTTTTCCCCA	ACGGGACGCT	CTACATCCGC	AACCTCGCGC	CCAAGGACAG	CGGGCGCTAT	6360
	GAGTGGGTGG	CCGCCAACCT	GGTAGGCTCC	GCGCGCAGGA	CGGTGCAAGT	GAACTGTCAG	6420
	CGTGCAGCAG	CCAAACGCGG	CATCACGGGC	ACCTCCCTCG	GGAGGACGGG	CGTCAGGTAC	6480
	GGAGGAACCC	TCAGCTGGGA	CTGCAGCGCC	TGGGGGAGCC	CCTGGGCGCG	CATCCTCTGG	6540
70	AGGCTGCGGT	CCAAGAGGAT	GATCGACGCG	CTCTTCAGTT	TTGATAGCAG	AATCAAGGTG	6600
	TTTGCCAAATG	GGACCTGCTG	GGTGAATCA	GTGACGGACA	AAGATGCCCG	AGATTACCTG	6660
	TGCGTAGCTC	GAAATAAGGT	TGGTGATGAC	TACGTGGTGC	TCAAAGTGGG	TGTGGTGTATG	6720
	AAACCGGCCA	AGATTGAACA	CAAGGAGGAG	AACGACCA	AAGTCTTCTA	CGGGGTGAC	6780
	CTGAAAGTGG	ACTGTGTGGC	CACCGGGCTT	CCCAATCCCG	AGATCTCTCTG	GAGCCTCCCA	6840
75	GACGGGAGTC	TGGTGAACTC	CTTCATGCG	TGCGATGACA	GCGGTGGAGC	CACCAAGCGC	6900
	TATGTGCTCT	TCACCAATGG	GACACTCTAC	TTTAAACGAG	TGGGGATGAG	GGAGGAGGGA	6960
	GACTACACCT	GCCTTGTCTG	AAATCAGGTC	GGGAAGGAGC	AGATGAGAGT	CAGAGTCAAG	7020
	GTGGTGACAG	CGCCCGCCAC	CATCCGGAAC	AAGACTTACT	TGGCGGTTCA	GGTGCCTTAT	7080
	GGAGACGTGG	TCAGTGTAGC	CTGTGAGGCC	AAAGGAGAAC	CCATGCCCAA	GGTGACTTGG	7140
80	TTGTCCCAA	CCACACAGGT	GATCCCCACC	TCCTCTGAGA	AGTATCAGAT	ATAECAGATG	7200
	GGCACTCTCC	TTATTCAGAA	AGCCAGCGGT	TCTGACAGCG	GCAACTACAC	CTGCTCTGTC	7260
	AGGACACAGC	CGGAGAGGGA	TAGGAGAGCG	GTGTGGATTG	ACGTCAACGT	CCAGCCACCT	7320
	AAGATCAACG	GTAAACCCAA	CCCATCACCC	ACCGTGGCGG	AGATGACAGC	CGGGGCGCAT	7380
	CGGAAACTGA	TTGACTGCAA	AGCTGAAGGC	ATCCCAACCC	CGAGGGTGT	ATGGGCTTTT	7440
	CCCGAGGGTG	TGGTTCTGCC	AGCTCCATAC	TATGGAAACC	GGATCACTGT	CCATGGCAAC	7500

5	GGTTCCCTGG	ACATCAGGAG	TTTGAGGAAG	AGCGAICTCG	TCCAGCTGGT	ATGCATGGCA	7560
	CGCAACGAGG	GAGGGGAGGC	GAGGTTGATC	GTGCAGCTCA	CTGTCTCTGA	GCCCATGGAG	7620
	AAACCCATCT	TCCACGACCC	GATCAGCGAG	AAGATCAGCG	CCATGGCGGG	CCACACCATC	7680
	AGCCTCAACT	GCTCTGCGCG	GGGGACCCCG	ACACCCAGCC	TGGTGTGGGT	CCTTCCCAAT	7740
	GGCACCGATC	TGCAGAGTGG	ACAGCAGCTG	CAGCGCTTCT	ACCACAAGGC	TGACGGCATG	7800
	CTACACATTA	GCGGTCTCTC	CTCGGTGGAC	GCTGGGGCCT	ACCCTGCTGT	GGCCCCCAAT	7860
	GCCGTGGGCC	ACACGGAGAG	GCTGGTCTCC	CTGAAGGTGG	GACTGAAGCC	AGAAGCAAAC	7920
	AAGCAGTATC	ATAACCTGGT	CAGCATCATC	AATGTTGAGA	CCCTGAAGCT	CCCCTGCACC	7980
10	CCTCCCCGGG	CTGGGCGAGG	ACGTTTCTCC	TGGAGCTGCC	CCAATGGCAT	GCATCTGGAG	8040
	GGCCCCCAAA	CCCTGGGACG	OGTTTCTCTT	CTGGACAATG	GCACCTTCAC	GGTTCGTGAG	8100
	GCTTCGGTGT	TTGACAGGGG	TACCTATGTA	TGCAGGATGG	AGACGGAGTA	CGGCCCTTCG	8160
	GTCACCAAGC	TCCCGGTGAT	TGTGATCGCC	TATCTCTCCC	GGATCACCAG	CGAGCCACCC	8220
	CGGTCATCT	ACACCCGGCC	CGGGAACACC	GTGAAACTGA	ACTGCATGGC	TATGGGGATT	8280
	CCCAAGCTG	ACATCAGGTG	GGAGTTACCG	GATAAGTCGC	ATCTGAAGGC	AGGGGTTTCA	8340
15	GCTCGTCTGT	ATTGGAACAG	ATTTCTTCA	CCCCAGGGAT	CACCTGACCAT	CCAGCATGCC	8400
	ACACAGAGAG	ATGCGCGCTT	CTACAAGTGC	ATGGCAAAAA	ACATTCTCGG	CAGTGAAGCT	8460
	AAACAACATT	AAACCAAGCT	CTTCTGAAAT	GTGGATTCCA	GAATGATGTC	TTAGGAACCT	8520
	ACACAAAGGC	GGGTTTGTGA	AGGGAAGCCA	GGTTGGGGAA	TAGGAGCTCT	TAAATAATGT	8580
20	GTCACAGTGC	ATGGTGGGCT	CTGGTGGGTT	TCAAGTTGAG	GTGATCTTGT	ATCTACAATT	8640
	GTGGGGAATA	GGAAGCAATG	CAGACACGAG	AAGGAGGGCT	CAGCCTTGCT	GAGACACTTT	8700
	CTTTTGTGTT	TACATCATGC	CAGGGGCTTC	ATTCAGGGTG	TCTGTGCTCT	GACTGCAATT	8760
	TTTCTTCTTT	TCCAAAGGCC	ACTCGACTGC	CTTCATAAGC	GTCATAGGTA	TATCTAGGGA	8820
	ACATTTCATC	AAATAAGGCC	ATAGACATGA	ACACACCTCT	ACTACCCCAT	TGAAGACGCA	8880
25	TACCTAGTAT	AACCTGCTGC	AGTTTTTACA	TGATAGACTT	TGTTCCAGAT	TGACAGTCA	8940
	TCCTTCAGTT	ATTTCCTCTG	TCACCTCAAA	ACTCCAGCTT	GCCCAATAAG	GATTTAGAAC	9000
	CAGAGTGACT	GATATATATA	TATATATTTT	AATTCAGAGT	TACATACATA	CAGCTACCAT	9060
	TTTATATGAA	AAAGAAATAA	CAITTTCTCC	TGGAACTCAC	TTTTATATA	ATGTTTTATA	9120
	TATATATTTT	TTCTTTTCAA	ATCAGACGAT	GAGACTAGAA	GGAGAAATAC	TTTCTGTCTT	9180
30	ATTAAATATA	ATAAATTTAT	GGTCTTTACA	AGACTTGGAT	ACATTACAGC	AGACATGGAA	9240
	ATATAATTTT	AAAAAATTTT	CTCCAAACCT	CCTTCAAAAT	CAGTCACCAC	TGTTATATTA	9300
	CCTTCTCCAG	GAACCTTCCA	GTGGGGGAAG	CTGCGATATT	AGATTTCCTT	GTATGCAAA	9360
	TTTITGTGAA	AGCTGTGCTT	CAGAGGAGGT	GAGAGGAGAG	GAAGGAGAAA	ACTGCATCAT	9420
	AACTTTACAG	AATTGAATCT	AGAGTCTTCC	CGAAAGGCC	CAGAAACTTC	TCTGCAGTAT	9480
35	CTGGCTGTGC	CATCTGCTGT	AAGTGGCTG	CTTCTTCCCC	AGCCATGAGT	CAGTTTGTGC	9540
	CCATGAATAA	TACAGACCTT	GTTATTTCCA	TGACTGCTTT	ACTGTATTTT	TAAGGTCAAT	9600
	ATACTGTACA	TTTGATAATA	AAATAATATT	CTCCCAAAAA	AAAAA		

Seq ID NO: 58 Protein sequence
Protein Accession #: NP_056234.1

40	1	11	21	31	41	51	
	MPKRAHWGAL	GVVLLILLMGH	PRVALACPHP	CACYVPSEVE	CTFRSLASVP	AGIARHVERI	60
45	NLGFNSIQAL	SBTSFAGLTK	LELLMIGHNE	IPSIPDGAIR	DLSSILQVFK	SYNKLAVITG	120
	QTLQLGSLNM	RIHIDHNKYE	FIHPQAFNGL	YSLRLHLHLS	NLLHLQLHPST	FSTFTFLDYF	180
	ELSTIRHLYL	AKNVMRTTLP	SMLRNMPLE	NLYLQGNPWT	CDCEMRWFLE	WDAKSRGILK	240
	CKKDKAYEBG	QLCAGCFSEK	KLYKHEIEKL	KDMTCLKPSI	ESPLRQNRSE	SIEEBEQBSE	300
	DGGSQLLILEK	FQLPQWSISL	NMTDEHGNMV	NLVCDIKKPM	DVYKIHINOT	DPPDIDINAT	360
50	VALDFPCPMT	RENYEKLWKL	IAYTSEVPVK	LHRELALSLD	PRVSYQYRQD	ADREBALYTG	420
	VRAQILAEPE	NVMQPSIDIQ	LNRRQSTAKK	VLSYVYTVQS	QTISTKQTRQ	ARGRSWVMIE	480
	PSGAVQBDQT	VLEGGFCQLS	CNVKASESPS	IFNVLPDGSY	LKAPMDPDPS	KFSILSSGWL	540
	RIKSMESPEGS	GLYQCIAQVR	DEMIRMVYRV	LVQSPSTQPA	EKDTVTIGKN	PGESVTLECN	600
	ALATPEAHLS	WILPNRRRIIN	DLANTSIVYM	LNGLTSLIPK	VQVSDSGYR	CUAVNQQCAD	660
55	HFTVGIIVTK	KSGSLPSKRG	REPQAKALSR	VREDIVDESG	SGSGMDEBNT	SRLHLHPKQD	720
	EVFLKTKIDA	INGDKAKKRG	RRKLKLWKHS	EKEPETNVAE	GRRVPESRRR	INMANKQINP	780
	ERNADILAKV	REKNLPKGTSE	VPPLIKITSP	PSLSLEVTFP	FPAVSFPASAS	EVQVTSABES	840
	SSADVPLIGE	KEEVLEITLSS	ASMGLEHNNH	GVILVESEVT	STPLEEVVDD	LSEKTESITS	900
	TSGDLKGTAA	PYLISKEVYEP	SEPTLHLDTV	YEKPTHESTA	TGWSAADVQ	SSPEPTSESE	960
60	EPFLDAVSLA	ESEPMQYDFD	DLETKSQPDE	DKMKEDTFAH	LTPTPTIWNV	DSSTSQLFED	1020
	STIGEPGVPG	QSHLQGLTIN	IHLVKSSLST	QDTLLIKKGM	KEMSQTLQGG	NMLBGDPSES	1080
	RSSSESQGES	KSTITLPSDTL	GIMSSMSFVK	KPAETTVGTL	LKDITTTVT	TPROKVPASS	1140
	TMSTHPSRRR	PMGRRRLRPN	KFRHRHKQTP	PTTFAPSETF	STOPTQAPDI	KISSQVESSL	1200
	VPTANVONTV	NTPKQLEMEK	NAEPTSKGTF	RRKRGKRPNK	HRYTPTSTVSS	RASGSKPSFS	1260
65	PENKGRNIVT	PSSETILLFR	TVSLKTEGTY	DSLDMYMTTR	KIYSSYPKVQ	ETLPVTVKPT	1320
	SDGKEIKDDV	ATNVDKHKSD	ILVTGESITN	AIPTSRSLVS	TMGEFKESSE	PVGFPPTPTW	1380
	NPSTTAQPGR	LQTDIPVTTT	GENLTDPLLL	KELEDVDFTS	EPLSSLTVST	PFGHEAGSS	1440
	TTLSSEIKVEV	ASSQAEITTL	DQDHLKITVA	ILLSETRPQN	HTPTAARMKE	PASSPSTIL	1500
	MSLGQTTTTK	PALPSRISQ	ASRDSKENVF	LNVCNPNPETE	ATTVNNEGTQ	HMSGPNELST	1560
70	PSSDRDAFNL	STYKLELEKQV	FGSRLPRGPF	DSQEQDGRVH	ASHQLTEVPA	KPILFATVR	1620
	LPEMSTQAS	RYFVTBQSER	HWINKPEITT	YPGALPENK	QFTTPRLSST	TIPLEHMSK	1680
	PSLPSKFTDR	RVDQPMNGYSK	VFGMMNIPFA	RNPVGRKPPSP	RIPHYENGR	PPFTNKLSP	1740
	PQLGVTRRRP	IPTSFAPVNR	ERKVIKPGSYN	RIHSHSTFHL	DFGPPADPLL	HTPQITGSPS	1800
	TNLQNLPMVS	STQSSISFIT	SSVQSSGSEH	QSSSKFPAGG	PPASKFWSLG	EKPQILTKSP	1860
75	QTVSVTAERT	TVFPCEATGK	PKPFTVWTKV	STGALMTFNT	RIQRFEVLKN	GTLVIRKVVQ	1920
	QDRGQIMCTA	SNLHGKDRMV	VLLSVTVQQP	QILASHYQDV	TUYLEDITAM	ECLANGTPAP	1980
	QISWIFPDOR	VMTQVSPVRS	RITLHENRTL	SIEKASFSDR	GVYKCVASNA	AGADSLAIRL	2040
	HVAALPPVIR	QSKLEINISLF	PGLSIHICT	AKAAPLPSVR	WVLGDGTQIR	PSQPLEHNL	2100
	VFPNGTLYIR	NLAPKDSGRY	ECVAANLVGS	ARRTVQLNVQ	RAANARITG	TSPRTDVRV	2160
80	GGTLKLDSCA	SGDFWPRILW	KLPSKRNIDA	LFSEDSRIK	FANGTLVVK	VIDKADAGD	2220
	CVARNKVDSD	YVVLKVDVVM	KPAIKERKEE	MHRKVFYGGD	LKVDVCATGL	PNPISWSLPL	2280
	DGSLVNSFMQ	SDDSGGRTKR	YVVFNNGTLY	FNEVGMREBG	DYTCFAENQV	GKDMENVRVK	2340
	VYTPATIRN	KTYLAVQVPY	GDVVTVACEA	KGEPMKVTW	LSPTNKVIPT	SSEKYQIYVD	2400
	GTLILYQAKR	SDSGNTYCLV	RNSAGEDRKT	VNIHVNQOPP	KINGNPNPIT	TVREIAAGGS	2460
	RKLIDCKABG	IPTRFVLNAP	PBGVVLPAFY	XGNRITVEGN	GSILDIRSLK	SDSVQLVCM	2520

RNEGGEARLI VQLTVLEPME KPIFHDPISE KITAMAGHTI SLNCSAAGTP TPSLVWVLPN 2580
 GTDLQSQOQL ORFYHKAIDGM LHSGLSSVD AGAYRCVARN AAGHTERLVS LKXVLKPEAN 2640
 KQYHNLVSLI NGETLKLPLCT PPGAGQGRFS WLEPNMHELE GPQTLGRVSL LDNGTLTVRE 2700
 ASVFDRTYV CRMETSYGPS VTSIPVIVIA YPPRITSEPT PVLYTRPGNT VKLNCMAMGI 2760
 PKADITWELF DKSHLKAGVQ ARLYGNRFLH PQGSLTIQHA TQRDAGFYKC MAKNILGSDS 2820
 KTTYIHVF

Seq ID NO: 59 DNA sequence
 Nucleic Acid Accession #: Bos sequence
 Coding sequence: 1..5001

1 11 21 31 41 51
 15 ATGCCAGGCA CAAACTAATC CGAACAGGC GCCCAGCAG ACTACAGAGT GATATTGAAG 60
 ACCTCTCAAG AGGACGAATT GGATGTACCT GACGACATCA GCGTCCGGGT TATGTCATCT 120
 CAGTCTGTGC TTGTGTCTGT GGTGGATCCY GTTCTGGAAA AACAGAGAA AGTTGTTCGA 180
 TCAAGACAGT ACACCGTGGC CTATCGAGAG AAGGGGGAAT TGGCCAGGTG GGATTATAAG 240
 CAGATCGCTA ACAGGCGTGT GCTGATTGAG AACCTGATT CAGACACTGT GTATGAATTT 300
 GCAGTCCGTA TTTCACAGGG TGAAGAGAT GGCAAATGGA GTACGTCACT CTTCCAAAGA 360
 20 ACACCAGAT CTGCCCTAC CACAGCTCC CACAACTTGA ACGTCTGGCC AGTCAATGGC 420
 AARCTACAG TTGTGCTGTC ATCTTGGGAT GCGCTACCA AGACTGAGGG GAAAGTGAAA 480
 GTCTGTCTGC TGGACACAGG ACTGTTTTC GTTCTCTCT TCCAACCATC TGCCAAATCA 540
 TTTCAAGATA CATCTTTTCA TACGCCCCG CTCTCAAAAC ATTGGAGCA AAGTCCCTCA 600
 25 CCAATCTTGG AGCACTACT TCTGCGCTG TGGATGGTCT GCAGCCTGGG GAACGCTATC 660
 TTTTCAAAAT CCGGGCCACA AACAGGAGAG GCGTGGGACC TCACTCCAAA GCGTTCATTG 720
 TCGCTATGCC AACAGAGATG CAGCTGTACC CAGAAGGATT TCACTGTGCT AGCTTACCTG 780
 ATCGATATCC AAACCAAAACA AGTTAATAAA GATCCACAAC TGGAAAGGAG TGTTTTGGGA 840
 CCATGTTTTC TTTTCTACTT CTTCACTATT ATGCTGGATA TTGGCGGCTT TTCTTCATT 900
 30 ATGTGCTATG AAGACCCANN TGTGTTCTCT TTGACAGGCA ATTCCTTAAA ATCTGTGTGA 960
 GCCAGTAAGG CCGATGTTCA GCAGAACACG GAGGACAAAT GGAACCCGGA AAAACCTGAG 1020
 CTTCTCTCAC CTTCTCCAG AGCTCCAGCT TCCTCCCAAC ACCCTCTGT GCTTGCTTCT 1080
 CCCCAGGGA GAAATGCCAA GGACCTTCTT CTGACTTGA AGAACAAAT ATTGGCTAAT 1140
 GGTGGGGCGC CCGAAAACC CCACTTCCG GCCAAGAGG CAGAGGAGCT GGATCTTCAG 1200
 35 TCAGACAGAA TCATCGGGGA GGAGGAGCTG GGTCCCGGG AGGACTCGCC CATGTCAACC 1260
 TCAGACACCC AAGGCCAGAA ACGGACCTG AGGCCGCCAA GTAGACAGG CCACTCGGTG 1320
 GTTGTCTCCG GCAGGACTGC AGTGAGGGCC CGGATGCCAG CCGTGCCTCG AAGGGAAGGC 1380
 GTAGATAAGC CTGCTTTTTC CTTGGCCAG CAGCCCGCC CAGGGGCGCC CCCCCTGGCT 1440
 TCGGCTCTCT CTGCCACCA CGGTCACACC CAGGGCACCT CTCTCGTCC TTCCCTGCCT 1500
 40 GCGAGCTTGA ATGACAAAGA CTGCTGGGAC TCAGACGAAG ATGAGCGCGC TGTGGGCTCC 1560
 CTCCACCCCA AGGGCGCCTT GCGCCAGCC CCGCCAGCCC TGTCCCGCAG CGCCAGTCC 1620
 CCGTCCAGCG TTCTCCGCGA CAGAAGCTCT GTCCACCCCG GCGCAAAAGC AGCCTCCCGC 1680
 GCGCGGAGGA CCCCOCATT CAGGGGCCGA GAGGAGATTT CCAGTGCCTC AGCCCCACCC 1740
 TCAGACTTTT CTCCACCCA TGGGGATCA TCTCGGCTGC TCGCCACCCA GCCACACTG 1800
 45 AGCTCTCCAC TTTCAAGGG CGGGAAGGAT GGTGAGGAGC CCGCAGCCAC CAACTCCAA 1860
 GCGCATCAC GGGTCCCTC GTCTCTCTC GTCTCTCTC ATCTCTCTC CAGGAGCAG 1920
 GTCTCTGAGG GAGCGGAGGC TTCTGATGAT GAAAGCCAG GTGACGCGA TAGGGAAGAC 1980
 GCGGGAAGGC AGGCGGAGGC CAGGCGCCAG ACCTGCGGG CCGGCGCTGC CTCTGACAC 2040
 50 TTCCATTTGC TCAGACAAA ACCCTTTGCT GCCAACGGGA GGTCTCCAG CAGGTTGAGC 2100
 ATTGGGCGGG GACCTCGGCT GCAGCCCTCC AGCTCCCCAC AGTGCAGTGT GCGCTCCCA 2160
 GCGCACCCCA GGGTCCCTC TCACTCTGAT TCCACCTTA AGCTTAGCTC AGGTATCAT 2220
 GGAGACGAGG AGGATGAGAA GCGGCTTCTT GCGACCGTTC TCAATGACCA CGTCCCTTCC 2280
 TCCTCCAGGC AGCCCATCTC CCGGGGCTGG GAGGACTTAA GGAGAGGCC GCAGAGAGGG 2340
 55 GCGAGCTTGC ATCGAGAGGA ACCCATCCCA GAGAACCCCA AATCCACAG GGCAGATACA 2400
 CATCTCAGG GCAAGTACTC CTCCCTGGCC TCCAGGGCTC AGGATGTTC AAGAGACACA 2460
 GAGCGGACA CCGGAGGTCA TTCTCCCAA GCACAGCCAG GGTCCACAGA CCGCCACGG 2520
 TCCCTGCTC GTCTCCCGC AGCAGGCTCA CAGCAGATC CAGTGTTC CAGAAGGATG 2580
 ACACCTCGCT GGGCCCCAGA ACAGCAGCCC CTTCTTCCG TCGCCAGCTC CAGCAGCAC 2640
 60 CCGGGAACCC CAGAGAGAGA CGGGGCTCG TCACTTCCG AGCCAGGCT CTCACTGACC 2700
 CAGGCTCGGC GCGCCCGCC CAGCTCGCAG GCGCGCTCC ACTCTCTC GGCCTTCC 2760
 ACGGCGAGCT GCGAGGGAT GCTCCCTCAG GCGCTCCAGA ACCAGGAGG GATGCTCCAG 2820
 GCGAGCTACG ACCAGACAG CACAGAAATC GAGGCGCAGG ATGTGCGGG CCGCGGCGAC 2880
 GCGCGCGCG CCAAGGAGGC AGCTGCGTCC CTTCCCAAGC ACCAGCAGT GAGTCTTCC 2940
 65 ACAGGCGCAG GGGCAGGTGG CAGCCACAGG TCCAGCGCG GACATGCGGC CTCCCGCGC 3000
 AGGCGCAGCC GACCGGCGG CCGCCAGTCC CCGCGCGGG TCCCGAGAG GGCAGGCGCG 3060
 GGGAGTCCG AGCTCTTTC CAAGCGGCCC CTGCTCTCCA AGTCCAGCA GTCGGTCTCA 3120
 GCGGAGGAGG AGGAGGAGGA GAGCGCGGG TTTTAAAG GCGGGAAGA AGACCTTCTG 3180
 70 TCTTCTCTG TGCAAAGTG GCGCTCTTCC TCCACTCCCA GGGCGGCAA AGACGCGAT 3240
 GGGAGCTCTG CCAAGGAAGA GAGGGAGGCT GCGATCGCG TCGCCCTCG CCGAGGGAGC 3300
 CTGGCTCTG TGAAGGAGC TCTCCGCCA CCGCAGGCA GCTCCCGAG GCGCTTCCAC 3360
 GTCCCTTCCG GACCGCGCC TCGCAGCGCT GCGACCGTGA GCGCGGTGCG GGCACCTCAC 3420
 75 CCGTGGCGCC GGTACACAC CCGCGCGCCV CCGCGCACT TCTCCACAC CCGATGCTG 3480
 TCTTGTGCGC AGAGGATGAT GCGATGCCAG TCCGTAACCT CTCTCTCCG ACAGCGTCC 3540
 AGACCTCTT ACAGCAAGG TTATAATGGC AGACCAATG TAGAAGGGA AGTCTTCTCT 3600
 GGTAGTATG GAAACCGAA TGGACAGAGA ATTATCAATG GCGCTCAAG AACAAAGTG 3660
 GTTGTGGACC TTGTGCTGG GTTAGTATG AATGACAGAG GAGGTACCT CCAAGATTCA 3720
 CATGGAATC CTCTTCTGAT TAACTAGGA GAGATGGTC GAACCATGT AGATCTGGA 3780
 80 GGGACCCCGG TGTGAGTCC TGACGGCTC CCACTCTTG GCGAGGGGG ACATGGCACA 3840
 CCTCTGGCCA ATGCCAAGA TAAGCCATT TTGAGTCTT GAGGAAGGC GCTGTGGG 3900
 TTGAGGTGCA TCAAAAAAC CACCATCCC CCACTACCTA CCACTGAGCC CACCACTACT 3960
 ACGAGGCCCC TGATCTGCT TACAACCCG AGGCCACCA CTGCCACCA CATGACGCCC 4020
 ACCACTACTA CAGCGCCCT GCTTACCCT ACACCGAGGC CCACTACTG CACCACTCC 4080
 GCGACGACCA CAGGCGTCC AACCAACCA GTCCGAACCA CTACGCGAC AACCACTACC 4140
 ACCACCCCA ACCCATCCC TCCATCCCC ACCGTCCCC CTGGAGCTT GGAACGCGAC 4200
 GACGATGATG GCAACTGAT AATGAGCTCC ATGGGATCC CAGAGTCTA CGCTGAAGA 4260

5
10
15
20
25
30

```

GATGAGTTCT CAGGCTTGGG GACTGACACT GCAGTACCTA CGGAAGAGGC CTACGTTATA 4320
TATGATGAAG ATTATGAATT TGAGACGTCA AGGCCACCAA CCACCACTGA GCCTTCGACC 4380
ACTGCTACCA CACCGAGGGT GATCCAGAGG GAAGGCGCCA TCAGTTCCTT TCCTGAAGAA 4440
GAATTTGATC TGCGTGGGAG GAAACGATTT GTTGCTCCTT ACCTGACGTA CCTAAATAAA 4500
GACCCATCAG CCCCGTGCTC TCTGACTGAT GCACGTGGATC ACTTCCAAGT GGACAGCCTG 4560
GATGAATACA TCCCAATGA OCTGAAGAAG AGTGATCTGC CTCCCAGCA TGCTCCCGC 4620
AACATCACCG TGGTGGCCGT GGAAGGTTGC CACTCATTTG TCATTGTGGA TTGGGACAAA 4680
GCCACCCAG GAGATTTGGT CACAGGTTAT TTGGTTTACA GTGCATCCTA TGAAGATTTC 4740
ATCAGGAACA AGTTTTCAC TCAAGCTTCA TCAGTAACTC ACTTGCCCAT TGAGAACCTA 4800
AAGCCCAACA CGAGGTATTA TTTTAAAGTG CAAGCACAAA ATCTCATGCG CTACGGACCT 4860
ATCAGCCCTT CGGTCTCATT TGTACCGAA TCAGATAATC CTCGTCTGTG TGTGAGGCC 4920
CCAGGCGGTG AGCTATCTG ATCCCATTCG CTTTCAAACA TGATCCGAGC TACACGGACT 4980
GCCATGGACG GCAATATGTA AAGCGCACGT GGTATCGAAA GTTCGTGGGA GTTGTTCCTT 5040
GTAATTCACT GAGGTATAAA ATCTACTTCA GTGACAACTT GAAAGATACA TTCTACAGCA 5100
TTGGAGACAG CTGGGAGACA GGTGAAGACC ATTGCCAATT TGTGGATTCA CACTTGATG 5160
GAAGAACAGG GCCTCAGTCC TATGTAGAAG CCTCCCTAC TATTCAGGCG TACTATCGCC 5220
AGTATCGTCA GGAGCCTGTC AGGTTTGGA ACATCGGCTT CGGAACCCCG TACTACTATG 5280
TGCGCTGGTA CGAGTGTGGG GTCTCCATCC CTGGAAGTGT GTATACAG GACCGTCATG 5340
CTGCAAGCTT GCCCTGCCCA GCCCCACCAA CTAAGTCGCA CTAGGGGCTG TGAGCAAGA 5400
CAGCCAGCAT GCTGAGCCCG GCTGCCCTAG GTGCCAGGAA GGTACAGAT GGACACTGGC 5460
CATTCGTGTC ATCTCAGTCT GGAACCTCAGT CCCACTTCTT GGCCTGGACA ATGAACAGGA 5520
TTCAGTTTGT CTGTTAACTT TGCTTCCTA CTTTTTTTGT TTGTTTGT TAAGCACATC 5580
CCAGAGACAT CAGAAACAG CACTGATTC AGTGTGATP CCCAGACTTT TTAGGCATGA 5640
AATTCGGACA CTTCACTATT TCCAGGAATA GCATATGCAC GCTGTTCTTG CTTCTAGGAA 5700
TGCTACATGC TTTCTGTTTT TCTCATTTTG GATTTCCTCA AAACCTAAGT AATTAAAGCT 5760
TCAGGTCCCT TTGTATGCGG TAGAAGGAA TTATTAAGAA CACCACCAA GAAATTAAT 5820
ATATCTACT TGAATTTAC TCTATGGACT TACCCACTGC TAGAATAAAT GTATCAATC 5880
TTATTTGTA ATCTCAATT TTGATATATA TATGATATATA TATCCACACT 5940
TGCTGCAAG AATATTGATT AATATTGCTA AATTTGTACT TGTTCACCAA AAAAAA 6000

```

Seq ID NO: 60 Protein sequence
Protein Accession #: Eos sequence

35
40
45
50
55
60
65

```

1 11 21 31 41 51
MFGTKLRTG APADYRVILK TSQEDLDLVP DDISVRVMS QSVLVSWVDP VLEKQKVVVA 60
SRQYTVRYRE KQELARWDYK QIANRRVLIE NLIPDTVYEF AVRISQGERD GKWSTSVFQR 120
TPESAPTTAP ENLNVWFVNG KPTVVAASWD ALPETEGKVK VCLLDITGLFS VSSFPQSAKS 180
FQNTFFETPR LSNHLEBSPS FILETILLPW WMVCSLGNAI FSKSGFQTGE AWDLTPKPSL 240
SLCQQBECST QMDFSLAYL IDIQTKQVWK DPQLEGSVPG PCFLEYFLTF MLDIGGFSFI 300
MCYEDPVSSL TGNLSKSVAA SZADVQWTE DNGKPEKPEP SSPSERAPAS SQHPSPVPSF 360
QGRNAKDLL DLKNKILANG GAPRKPQLRA KKAELDLQS TEITGEELG SREDSPMSPS 420
DTQDQKRLR PPSRHHGSHV APGRATVRAR MPALPREGV DKGPFSLATQ PRPGAPPSAS 480
ASPAHHASTQ GTSRHPSLFA SLNDMDLVDS DEDERAVGSL HPKGAFAQPR PALSPSRQSP 540
SSVLDRSSV HPGAKPASTA RRTPHSGAAE EDSSASAPP RLSPPHGSS RLLPTQPHLS 600
SPLSKGKDG EDAPATNSNA PSRTMSSSV SSSLSSRIQV SEGAEASDGE SHGDGREDG 660
GRQAEATAQT LKRAPASGEF HLLREKPFPA NGRSPSRFSI QEGPRLQPS SPQSTVPSRA 720
HPRVPSHSDS HPKLSGLIIG DEEDEKPLFA TVVNDHVPSS SRQPSIRGWE DLRRSPQSGA 780
SLHRKEE LPE NPKSTGADTH PQGYSSLAS KAQDVQSTD ADTEGHSFKA QPGSTDRHAS 840
PARPEAARSQ QHPSVRRMT PGRAPQQPP PVVATSHHP GPQERDAGRS PSQRLSLTQ 900
AGRPRPTBQC RSHSSSDPYT ASSRGMLETA LQXQDEDAQ SYDDSTVEVE AQDVRAPARA 960
ARAKBAASL PKKQVSEPT GAGAGDERS QRGHAAAPAR PRRPGGPQSR ARVPSRAAFG 1020
KSEPPSKRPL SSKSQSVSA EDDEEDAGF PKGKEDLLS SSVKWPSSS TPRGKDADG 1080
SLAKEERER LALAPRGSL APVKRPLPPP PGSSPRASHV PSRPPPSAA TVSEVAGTHP 1140
WPRYTTTRAP GFTTTPMLL LRQRMHARF RNPLSRQPAR PSYRQYNGR ENVEGKVLPG 1200
SNKPFNQRI INGFQTKHV VDLDRGLVLN AEGRYLDSE CNPLRIKLG DGRITVILEG 1260
TPVSPDGLP LFGQGRGTF LANAQDRPIL SLGKPLVGL EVIKKTHPP TTTMQPTTT 1320
TPLPTTTTPR PTTATMQPT TTTTLPITP PRPTATTTR TTTRPTTV RTTTRTTTT 1380
TPKPTTPIPT CPPGTLERHD DDGNLIMSSN GIPECYAEED EFGLETDIA VPTERAYVIY 1440
DEDYEFETSR PPTTTEPST ATTPEVPEE GAISFPSEE FDLAQRKRFV APVVTYLNKD 1500
PSAPCSLIDA LDHFPVDSLD ELIPNDLKS DLPQHAPRN ITVVAVEGCK SFVIVDWDKA 1560
TPGDLVGYL VYSASYEDFI RNKFTQASS VTHLPFLENK ENTRYFKVQ AQNPHGYGPI 1620
SPSVSFVTS DNPLLVVRP GGLBGSHSL SMMPATRTA MDGMM

```

Seq ID NO: 61 DNA sequence
Nucleic Acid Accession #: NM_022743
Coding sequence: 128..1237

70
75
80

```

1 11 21 31 41 51
GTGGATTTTA GAGATACCTC CCTCCTTCT GCTCAGCTGC CTTCAGTAA TTAAACTCTT 60
TCTCTGCTGC AACACCCCTA CTGTTCTCGT TGTATTGGCT TTTCTGGGCA GCAGGAAGGA 120
AAGCTGTAIG CGATGCTCTC AGTGCCGCGT CGCCAAATAC TGTAAGTCTA AGTGTGAGAA 180
AAAAGCTTGG CCAAGACACA AGCGGGAATG CAAATGCCTT AAAAGCTGCA AACCCAGATA 240
TCTCTCAGAC TCGGTTGAGC TTCTTGGCAG AGTTGTCTTC AAACCTATGG ATGGAGCACC 300
TTCAGATACA GAGAAGCTTT ACTCATTTTA TGATCTGGAG TCAATATTA ACNAACTGAC 360
TGAAGATAG AAAGAGGGCC TCAGGCAACT CGTAATGACA TTTCAACATT TCATGAGAGA 420
AGAAATACAG GATCCCTCTC AGCTGCCACC TGCCCTTGAC CTTTTGAG CCTTTGAAA 480
AGTATCTGCG AACCTCTTCA CCATCTGTAA TCGGAGATG CAGGAAGTTG GTGTTGGCCT 540
ATATCCAGT ATCTCTTTCG TCAATCACAG CTGTGACCCC AACTGTTGCA TTGTGTTCAA 600
TGGGCCACAC CTCCTTACTGC GAGCAGTCCG AGACATCGAG GTGGGAGAGG AGCTCACCAT 660
CTGCTACCIG GATATGCTGA TGACCACTGA GGAGCGCCGG AAGCAGCTGA GGGACCAGTA 720
CTGCTTTGAA TGTGACTGTT TCGTGTGCCA AACCCAGGAC AAGGATGCTG ATATGCTAAC 780

```

5 TGGTGATGAG CAAGTATOGA AGGAAAGTTCA AGAATCCCTG AAAAAAATTG AAGAACTGAA 940
 GGCACACTGG AAGTGGGAGC AGGTTCCTGG CATGTGCCAG GCGATCATAA GCAGCAATTC 960
 TGAACGGCTT CCCGATATCA ACATCTACCA GCTGAAGGTG CTCGACTGCG CCATGGATGC 960
 CTGCATCAAC CTCGGCTCTG TGGAGGAAGC CTGTCTCTAT GGTACTCGGA CCATGGAGCC 1020
 ATACAGGATT TTTTCCAG GAAGCCATCC CGTCAGAGGG GTTCAAGTGA TGAAGTTGG 1080
 CAAACTGCAG CTACATCAAG GCATGTTCC CCAAGCAATG AAGAATCTGA GACTGGCTTT 1140
 TGATATTATG AGATGCACAC ATGGCAGAGA ACACAGCCTG ATTGAAGATT TGATTCTACT 1200
 TTAGAAGAA TGGACGCCA ACATCAGAGC ATCTAAGGG AACGCAGTCA GAGGGAATA 1260
 CGCGCTGTGT CTGTGTGAA TGCCTTATTG AGGTCAACAC CTCTATGCTT TGTAGCTGT 1320
 10 GTGAACCTCT CTATTGAA ATTCTGTTCC GTGTTGTGT AGGTAAATAA AGGCAGACAT 1380
 GGTTCGAAA CCACAGAAT CATTAGTTGT AGAGAAGCAC GATTATAATA AATTCAAAC 1440
 ATTTGGTTGA GATGCCAAA AAAAAAAAAA AAAAAA

15 Seq ID NO: 62 Protein sequence
 Protein Accession #: NP_073580

1 11 21 31 41 51
 20 MRCSQCRVAK YCSAKCKKKA WPDEKRECKC LKSCCKPRYPF DSVRLILGRVV FKLMGAPSE 60
 SERLYSPYDL ESNINKLTED KKEBLRLQVM TFQHFMRBEI QDASQLPPAF DLFEAPAKVI 120
 CNSFTICNAE MQEVGVGLYP SISLINESCD FNCSTVFNPF HLLLRVRDI EVGSELTICY 180
 LDMLMTSEER RKQLRDQYCF ECDCFRCQQT DKDADMLTGD EQVWKEVQES LKKIEELKAH 240
 WKWEQVLAMC QAIISNSNER LPDINIYQLK VLDCAMDACI NLGLLEELF YGRTMEPYR 300
 25 IFFPGSHFVR GVQVMKVGLK QLHQGMFPQA MNKLRLAFDI MRVTHGREHS LIEDLILLLE 360
 ECDANIRAS

30 Seq ID NO: 63 DNA sequence
 Nucleic Acid Accession #: NM_003014.2
 Coding sequence: 238..648

1 11 21 31 41 51
 35 GCGGGTTCTG CGCCCCGAG GCTGAGAGCT GCGGCTGCTC GTGCCCTGTG TGCCAGACGG 60
 CGGAGCTCGG CGGCGGAGCC CCGCGGCCCC GCTTTGCTGC CGACTGGAGT TTGGGGGAAG 120
 AAACCTCTCT GCGCCCCAGA AGATTCTTTC CTCGCGGAAG GGACACCGAA AGATGAGGGT 180
 GGCAGGAAGA GAAGCGGCTT TCTGTCTGCC GGGTTCGAG CGCGAGAGGG CAGTGCCATG 240
 TTCTCTCCA TCTAGTGGC GCTGTGCTTG TGGCTGCACC TGGCTGTGG CGTGCCTGGC 300
 GCGCCCTGGG AGGCGGTGGC CATCCCTATG TGCCGGCACA TGCCCTGGAA CATCACGGG 360
 ATGCCCAACC ACCTGCACCA CAGCACGCGAG GAGAACGCCA TCTTGCCCAT CGAGCAGTAC 420
 GAGGAGCTGG TGGAGCTGAA CTGCAGCGCC GTGCTGCGCT TCTTCTCTG TGCCATGTAC 480
 GCGCCATTG GCACCTTGA GTTCTGCTGC GACCTATCA AGCGGTGCAA GTCCGTGTGC 540
 CAACGCGGCG GCGACGACTG CGAGCCCTTC ATGAAGATGT ACACCCACAG CTGGCCCGAA 600
 AGCCTGGCCT GCGACAGACT GCCGTCTAT GACCGTGGGG TGTGCAATTC GCTGAAGCC 660
 ATCGTCAAGG ACCTCCCGGA GGATGTTAAG TGGATAGACA TCACACCGA CATGATGTA 720
 45 CAGGAAAGCG CTCTGATGT TGACTGTAAA CGCTAAGCC CGGATCGGTG CAGGTGTAAA 780
 AAGGTGAAGC CACTTTGGC AACGTATCTC AGCAAAACT ACAGCTATGT TATTCATGCC 840
 AAAATAAAG CTGTGCAGAG GAGTGGCTGC AATGAGGTCA CAACGTTGGT GGATGTAAA 900
 GAGATCTTCA AGTCTCATC ACCCATCCCT CGAACTCAAG TCCCGCTCAT TACAAATTCT 960
 TCTTGCCAGT GCTCCACACT CCTGCCCAT CAAGATGTTT TCATCATGTG TTACGAGTGG 1020
 50 CGTTCAAGGA TGATGTTCTT TGAAATTGCG TTAGTTGAAA AATGGAGAGA TCAGCTTAGT 1080
 AAAAGATCCA TACAGTGGGA AGAGAGGCTG CAGGAACAGC GGAGAACAGT TCAGGACAA 1140
 AAGAAACAG CGGGGCGCAC CAGTGTAGT AATCCGCCA AACCAAGGG AAAGCCTCCT 1200
 GCTCCCAAC CAGCCAGTCC CAAGAAGAAC ATTAAGACTA GGAGTGCCTA GAAGAGACA 1260
 AACCCGAAA GAGTGTGAGC TAACTAGTTT CCAAGCGGA GACTTCGAC TTCTTACAG 1320
 55 GATGAGGCTG GGCATTGCCCT GGGACAGCCT ATGTAAGGCC ATGTGCCCTC TCCCTAACA 1380
 ACTCACTGCA GTGCTCTTCA TAGACACATC TTGCAGCAT TTTCTAAGS CTATGCTTCA 1440
 GTTTTCTTGT GTAGCCATC ACAGCCATA GTGGTAGGTT TGCCCTTGG TACAGAAGT 1500
 GAGTTAAAGC TGGTGGAAA GGCTTATGCG ATTGCAATTA GAGTAACCTG TGTGCATACT 1560
 CTAGAAGAGT AGGGAATAA ATGCTTGTTA CAATTGAGC TAATATGTGC ATTGTAAAAT 1620
 60 AAATGCCATA TTTCAACAA AACACGTAAT TTTTTCAG TATGTTTAT TACCTTTTGA 1680
 TATCTGTGT TGCATGTTA GTGATGTTT AAAATGTGAT GAAATATAA TGTTTTAA 1740
 AAGGAACAGT AGTTGCAATG ATGTTAAAAG ATCTTATGT GTTTATGGTC TGCAGAGGA 1800
 TTTTGTGAT GAAAGGGGAT TTTTGAAGA ATTAGAGAAG TAGCATATGG AAAATTATA 1860
 TGTGTTTTT TACCAATGAC TTCAGTTTCT GTTTTAGCT AGAAACTTAA AAACAAAAT 1920
 65 AATAATAAG AAAATAAAT AAAAAGGAGA GGCAGACAAT GTCTGGATT CTGTTTTT 1980
 GTTACCTGAT TTCCATGATC ATGATGCTTC TTGTCAACAC CCTCTAAGC AGCACCAGAA 2040
 ACAGTAGGTT TGCTGTACC ATTAGGAGTT AGGTACTAAT TAGTTGGCTA ATGCTCAAGT 2100
 ATTTTATACC CACAAGAGAG GTATGTCATC CATCTTACTT CCCAGGACAT CCACCTGAG 2160
 AATAATTGA CAGCTTAAA AATGGCCTTC ATGTGAGTGC CAATTTTGT TTTTCTCAT 2220
 70 TTAATATT TCTTGGCTA AATACATGTG AGAGGAGTTA AATATAATG TACAGAGAG 2280
 AAAGTTAGT TCCACCTCTG AATGAGAAT TACTTGACAG TTGGGATAG TTAATCAGAA 2340
 AAAAAGACT ATTTGCAGC ATTTTATCA CAAATTCAT AATTGTGAC AATTGGAGC 2400
 ATTTATTTTA AAAACAATT TTATTGGCCT TTGTCTACA CAGTAAGCAT GTATTTTATA 2460
 AGGCATTCAA TAATGCACA ACGCCCAAG GAAATAAAT CCTATCTAAT CCTACTCTCC 2520
 75 ACTACAGAG TGAATCACT ATTAGTATT TGGCATATTA TTCTCCAGGT GTTGTCTTAT 2580
 GCACATTAA AATGATTGA ACAATAAAA CTAGGAOCT GTATACATGT GTTTCATAAC 2640
 CAGCTTCTT TGCTTGGCCC TTTATTGAGA TAAGTTTTC TGTCAGAAA GCAGAACCA 2700
 TCTCATTTCT AACAGCTGTG TTATATTCCT TAGTATGCAT TACTCAACAA ACTGTGTGTC 2760
 TATTGGATAC TTAGTGTGTT TCTTCACTGA CAATACAGAA TAAACATCTC ACOGGAATTC

80 Seq ID NO: 64 Protein sequence
 Protein Accession #: NP_003005.1

1 11 21 31 41 51

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

Seq ID NO: 65 DNA sequence
 Nucleic Acid Accession #: BC010423
 Coding sequence: 248..1780

1	11	21	31	41	51	
CACAGCGTGG	GAAGCAGCTC	TGGGGGAGCT	CGGAGCTCCC	GATCACGCTC	TCCTGGGGGT	60
AGCTACGGCT	GGGTGTGTAG	AACGGGGCCG	GGGCTGGGGC	TGGGTCCCTC	AGTGGAGACC	120
CAAGTGGGAG	AGGCAAGAA	TCTGCAGCTT	CCTGCCCTTC	GGGTCACTTC	CTTATTCAAG	180
CTGTCAGCCG	GCTCCACAGG	AGATCTCGGT	GGAACCTCAG	AAACGCTGGG	CAGTCTGCCT	240
TTCAACCATG	CCCTGTGTCC	TGGGAGCCGA	GATGTGGGGG	CCTGAGGCTC	GGCTGTGCTC	300
GCTGTCTACT	CTGGCATCAT	TTACAGGCCG	GTGCCCCGCG	GGTGAAGCTGG	AGACCTCAGA	360
CGTGTAACT	GTGTGTCTGG	GCCAGGACGC	AAAACTGCCC	TGCTTCTACC	GAGGGGACTC	420
CGGCGAGCAA	GTGGGGCAAG	TGGCATGGGC	TGGGGTGGAC	GCGGGCGAAG	GCGCCAGGA	480
ACTAGGCGTA	CTGCATCCA	AATACGGGCT	TCATGTGAGC	CGGCTTTACG	AGGGCCGCTT	540
GGAGCAGCCG	CCGCCCCCAC	GCAACCCCTC	GGACGGCTCA	GTGCTCCTGC	GCAACGCACT	600
CGAGGCGGAT	GAGGCGAGT	ACGAGTCCCG	GGTCAGCACC	TTCCCCGCGC	GCAGCTTCCA	660
GGCGGGGCTG	CGGCTCCGAG	TGCTGGTGCC	TCCCCGCCCC	TCACTGAATC	CTGCTCCAGC	720
ACTAGAAGAG	GCCCAAGGCC	TGACCTTGCC	AGCCTCCTGC	ACAGCTGAGG	GCAGCCGAGC	780
CCCCAGCGTG	ACCTGGGACA	CGGAGGTCAA	AGGCACAACG	TCCAGCCGCT	CCTTCAAGCA	840
CTCCCGCTCT	GCTGCCGCTA	CCTCAGAGTT	CCACTTGGTG	CCTAGCCGCA	GCATGAATGG	900
CGAGCCACTG	ACTTGTGTGG	TGTCCCATCC	TGGCTGCTCT	CAGGACCAAA	GGATCACCCA	960
CATCTCCAC	GTGTCTTCTC	TTGCTGAGGC	CTCTGTGAGG	GGCCTTGAAG	ACCAAAATCT	1020
GTGGCAGATT	GGCAGAGAAG	GAGCTATGCT	CAAGTGCCCT	AGTGAAGGGC	AGCCCCCTTC	1080
CTCATACAAC	TGGACACGGC	TGGATGGGCC	TCTGCCAGT	GGGGTACGAG	TGGATGGGGA	1140
CACTTTGGGC	TTTCCCCCAC	TGACCACTGA	GCACAGCCGC	ATCTACGCTC	GCCATGTCTAG	1200
CAATGAATTG	TCCTCAAGGG	ATTCTCAGGT	CACGTGGGAT	GTCTTTGACC	CCCAGGAAGA	1260
CTCTGGGAAG	CAGGTGGACC	TAGTGTACGC	CTCGTGGTGT	GTGGTGGGTT	TGATCGCCGC	1320
ACTCTTGTTC	TGCTTCTTGG	TGGTGGTGGT	GGTGTCTATG	TCCCGATACC	ATCGGGGCAA	1380
GGCCAGCGAG	ATGACCCAGA	AATATGAGGA	GGAGCTGACC	CTGACCAAGG	AGAACTCCAT	1440
CCGAGGCTTG	CATCCCCATC	ACACGGACCC	CAGGAGCCAG	CCGAGGGAGA	GTGTAGGGCT	1500
GAGAGCCGAG	GCCCCACCTG	ATAGTCTCAA	GGACAACAGT	AGCTGCTCTG	TGATGAGTGA	1560
AGAGCCCGAG	GSCCCAGATT	ACTCCAGCTC	GACCAAGGTT	AGGGAGATAG	AAACACAGAC	1620
TGAATGCTG	TCTCCAGGCT	CTGGGCGGGC	CGAGGAGCAG	GAAGATCAGG	ATGAAGGCAT	1680
CAACAGGGCC	ATGACACATT	TTGTTCAGGA	GAATGGGACC	CTACGGGCCA	AGCCCAAGGG	1740
CAATGGCATC	TACATCAATG	GGCGGGGACA	CCTGGTCTGA	CCGAGGCTTG	CCTCCCTTCC	1800
CTAGGCGCTG	CTCCTTCTGT	TGACATGGGA	GATTTTAGCT	CATCTTGGGG	GGCTCCTTAA	1860
ACACCCCATG	TTCTTGCGGA	AGATGCTCCC	CATCCACTG	ACTGCTTGAC	CTTTACCTCC	1920
AACCCCTCTG	TTATCGGGA	GGGCTCCACC	AATTGAGTCT	CTCCCACTAT	GCATGCAGGT	1980
CACGTGTGTG	GTGCATGTGT	GCCGTGTGTA	GTGTTGACTG	ACTGTGTGTG	TGTGGAGGGG	2040
TGACTGTCCG	TGGAGGGGTG	ACTGTGTCCG	TGGTGTGTAT	TATGCTGTCA	TATCAGAGTC	2100
AAGTGAATG	TGGTGTATGT	GCCACGGGAT	TTGAGTGGTT	GCGTGGGCAA	CACGTCTCAG	2160
GTGTGGCGTG	TGTGTCTATG	GGCTGTGTGT	GACCTCTGCC	TGAAAAGACA	GGTATTTTCT	2220
CAGACCCGAG	AGCAGTATTA	ATGATGCAGA	GGTTGGAGGA	GAGAGCTGGA	GACTGTGGCT	2280
CAGACCCAGG	TGTGCGGGCA	TAGCTGGAGC	TGGAATCTGC	CTCCGGTGTG	AGGGAACTCT	2340
TCTCTACCA	CTTCCGAGCC	ATGGGGGCAA	GTGTGAAGCA	GCCAGTCCCT	GGGTCAAGCA	2400
GAGGCTTGAA	GGTTTACAGA	AGCCCTCTGC	CCTCTGGTGG	CCTCTGGGCC	TGCTCATGAT	2460
ACATATTTTC	TGTAAATATA	CATGCGCCCG	GAGCTTCTTG	CAGGAATACT	GCTCCGAATC	2520
ACTTTTAAAT	TTTTTCTTTT	TTTTTCTTTT	CCCTTTCATC	TAGTTGTATT	TTTTTTTAT	2580
TTTTTTTAT	ATTTTATTTT	AGAGTTTGAG	TCCAGCCTGG	ACGATATAGC	CAGACCTCTG	2640
CTGTAAAAAA	ACCAAAACCC	AAAAAAAATA	AAAAAAAATA			

Seq ID NO: 66 Protein sequence
 Protein Accession #: AAI10423

65
 70
 75
 80

Seq ID NO: 67 DNA sequence
 Nucleic Acid Accession #: NM_001203
 Coding sequence: 274..1782

1	11	21	31	41	51	
CGCGGGGGCG	GGAGTCGGCG	GGGCTTCGCG	GGACGGGGCG	AGTGGGGAGA	CGCGGGGGCT	60
GAGGAGCGCG	GAGCCGGGAG	CGCAGCGCGG	GGGTGGAGTT	CAGCCTACTC	TTTCTTAGAT	120
GTGAAGGAA	AGGAAGATCA	TTTCATGCCT	TGTTGATAAA	GGTTCAGACT	TCTGCTGATT	180

5
10
15
20
25
30

```

CATAACCAIT TGGCTCTGAG CTATGACAAAG AGAGGAAACA AAAAGTTAAA CTTACAAGCC 240
TGCCATAAGT GAGAAAGCAA CTTCCTTGAT AACATGCTTT TGGCAAGTGC AGGAAAATTA 300
AATGTGGGCA CCAAGAAAGA GGATGGTGAG AGTACAGCCC CCACCCCCCG TCCAAAGGTC 360
TTGCTTTGTA AATGCCACCA CCATTGTCCA GAAGACTCAG TCAACAATAT TTGCAGCACA 420
GACGGATATT GTTCAAGAT GATAGAAGAG GATGACTCTG GGTGCCCIGT GGTCACTTCT 480
GGTTGCCTAG GACTAGAAGG CTCAGATTTT CAGTGTCCGG ACACCTCCAT TCCTCATCAA 540
AGAAGATCAA TTGAATGCTG CACAGAAAGG AACGAATGTA ATAAAGACCT ACACCTTACA 600
CTGCCCTCAT TGAAAAACAG AGATTTTGTT GATGGACCTA TACACCACAG GGCTTTACTT 660
ATATCTGTGA CTGTCTGTAG TTGCTCTCTG GTCCTTATCA TATTATTTTG TTACTTCCGG 720
TATAAAGAC AGAAACCCAG ACCTCGATAC AGCATTGGGT TAGAACAGGA TGAAACTTAC 780
ATTCTCTCTG GAGAAATCCCT GAGAGACTTA ATTGAGCAGT CTCAGAGCTC AGGAAGTGGA 840
TCAGGCTCC CTCTGCTGGT CCAAGGAGCT ATAGCTAAGC AGATTTCAGT GGTGAACAG 900
ATTGGAAGAG GTCCCTATGG GGAAGTTTGG ATGGGAAGT GCGGTGGCGA AAAGTAGCT 960
GTGAAGATGT TCTTACCAC AGAGGAAGCC AGCTGGITCA GAGAGACAGA AATATATCAG 1020
ACAGTGTGA TGAGGCATGA AAACATTTTG GGTTCATTG CTGCAGATAT CAAAGGGACA 1080
GGGTCTCGGA CCCAGTTGTA CCTAATCACA GACTATCATG AAAATGGTTC CCTTATATGAT 1140
TATCTGAAGT CCACCAACCT AGACGCTAAA TCAATGCTGA AGTTAGCCTA CTCTTCTGTC 1200
AGTGGCTTAT GTCAATTACA CACAGAAATC TTTAGTACTC AAGGCAAAAC AGCAATTGCC 1260
CATCGAGATC TGAAAGTAA AAACATCTCT GTGAAGAAA ATGGAACCTG CTGTATTGCT 1320
GACCTGGGCC TGGCTGTTAA ATTTATTAGT GATACAAATG AAGTTGACAT ACCACCTAAC 1380
ACTGAGATTG GCACCAACAG CTATATGCCCT CCAGAAGTGT TGGACGAGAG CTGGAACAGA 1440
AATCACITCC AGTCTTACAT CATGGCTGAC ATGTATAGTT TTGGCCTCAT CCTTTGGGAG 1500
GTTGCTAGAG GATGTGTATC AGGAGGTATA GTGGAAGAAT ACCAGCTTCC TTATCATGAC 1560
CTAGTGCCCA GTGACCCCTC TTATGAGGAC ATGAGGGAGA TTGTGTGCAT CAAGAAGTAA 1620
CGCCCTCTGA TCCCAACCG GTGGAGCAGT GATGAGTGTG TAAGGCAGAT GGGAAACTC 1680
ATGACAGAAAT GCTGGGCTCA CAATCCTGCA TCAAGGCTGA CAGCCCTGCG GGTAAAGAAA 1740
ACACTTGCCA AATGTTCAGA GTCCAGGAC ATTAACCTCT GATAGGAGAG GAAAAGTAG 1800
CATCTCTGCA GAAGCCCAAC AGGTACTCTT CTGTTTGTGG GCAGAGCAGA AGACATCAA 1860
TAAGCATCCA CAGTACAAGC CTGAACATC GTCCGTGCTC CAGTGGGTT CAGACCTCAC 1920
CTTTCAGGGA GCGACCTGGG CAAAGACAGA GAAGTCCCA GAAGGAGAGA TTGATCCGTG 1980
TCTGTTTGA GCGGAGAAA CCGTTGGGA ACTTGTTCAT GATATGATGC AT

```

Seq ID NO: 68 Protein sequence

Protein Accession #: NP_001194

35
40
45

```

1 11 21 31 41 51
MLRSAGKLN VETKKEDGES TAPTPRPKVL RCKCRHHCPD DSVNNICSTD GYCFTMIRED 60
DSGLFVVTSQ CLIGEGSDPQ CRDTPFIPHQ RSIEOCTERN ECKNOLHPTL PPLKNRDFVD 120
GPIHRAALLI SVTVCSLLLV LILFCYFHY KRQETRPFRYS IGLBQDETYI PPGESLRDLI 180
EQSQSSGSGS GLPLLVQRTI AKQIQMVKQI GKGRYGEVWM GKWRGKVVAV KVFETTESAS 240
WFKETRIYQT VLMRHHLLG FIAADIKGTG SWTQLYLITD YHENGSLYDY LKSTTLDAXS 300
MLKLAYSBSVS GLCLHHTFIF STGKPAIAH RDLKSKNLIK KNGTCCCIAD LGLAVKFISD 360
TNEVDIPFNT RVGTGKRVMP EVLDBSLNRM HFQSYIMADM YSFLILMEV ARRCVSGGIV 420
EYQLFPHDL VPSDFSYEDM REIVCIKKLR PSEPNRNSD ECLRQNGELM TECAWHNEAS 480
RLTALRVKKT LAKMSBSQDI KL

```

Seq ID NO: 69 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 166..1737

50
55
60
65
70
75
80

```

1 11 21 31 41 51
TTGGGGGTTT ATTCTCTTCC CTCTAACTT GACAGGGTCT TGCCTCTGCA TTCAGGCAAG 60
AGTGCAGTAG TGTACTGACC TCTTACTGCC GCCTCAAGCT TCCAGCCTCA ACTCAAGCAA 120
TCTCCCAACC TCAGCCACCC AAGTGGCTGG GACTACAGAT TAAGAATGAC CCAAAATAAA 180
TTAAAGCTTT GTTCCAAAGC CAATGTGTAT ACTGAAGTGC CTGATGGAGG ATGGGGCTGG 240
GCGGTAGCTG TTTCATTTT CTTCGTTGAA GTCTTCACTC ACGGCATCAT CAAGACATTT 300
GGTGTCTTCT TTAATGACIT AATGGACAGT TTTAATGAAT CCAATAGCAG GATCTCATGG 360
ATAATCTCAA TCTGTGTGTT TGTCTTAACA TTTTCAGCTC CCTCGCCAC AGTCTGAGC 420
AATCGTTTGG GACACCGTCT GGTAGTGATG TTGGGGGGGC TACTTGTGAG CACCGGGATG 480
GTGGCGGCGT CCTTCTCACA AGAGGTTTCT CATATGTAGC TCGCCATCGG CATCATCTCT 540
GGTCTGGGAT ACTGCTTTAG TTTTCTCCCA ACTGTAAACA TCCTATCACA ATATTTTGGC 600
AAAAGACGTT CCATAGTCAC TGCAGTTGCT TCACAGGAG AATGTTTCGC TGTGTTTGCT 660
TTGCGACCCG CAATCATGGC TCTGAAGGAG CGCATTGGCT GGAGATACAG CCTCCTCTCT 720
GTGGGCTCAC TACAGTTAAA CATTGTCTATC TTGGAGCAC TGCTCAGACC CATCTTTATC 780
AGAGGACCCG CGTCACCGAA AATAGTCTATC CAGGAAAATC GGAAGAGAGC GCAATATATG 840
CTTGAAAATG AGAAAACAG AACCTCAATA GACTCCATTG ACTCAGGAGT AGAACTAACT 900
ACCTCACCTA AAAATGTGCC TACTCACACT AACCTGGAAC TGGAGCCGAA GGCOSACATG 960
CAGCAGGTCC TGGTGAAGAC CAGCCCCAGG CCAAGCGAAA AGAAGCCGCC GCTATTAGAT 1020
TTCTCCATTI TGAAAGAGAA AAGTTTATT TGTATGTCAT TATTTGGTCT CTTTGAACA 1080
CTGGGATTTT TTGCACTTCT CTGTGATATC ATTCTCTGCG GCATTAGTCT GGGCATTGAC 1140
CAGGACCGCG CTGCTTTTAT ATTATCTACG ATGGCCATTG CAGAAGTTTT CGGAAGGATC 1200
GGAGCTGGTT TTGTCTCTAA CAGGAGAGCC ATTCTGTAAGA TTTACATTGA GCTCATCTGC 1260
GTCTCTTAT TGACTGTGTC TCTGTTTGGC TTTACTTTTG CTACTGAATT CTGGGTCTCA 1320
ATGTCATGCA GCATATTTT TTGGTTTATG GTTGAACAA TAGGAGGGAC TCACATTTCA 1380
CTGCTTCTG AGGAIGATGT CGTGGGCATT GAGAGATGT CTCTGACAG TGGGGTCTAC 1440
ATCTTCATTC AGAGCTGGCT AGGACCGCCC TTGCAGTTTT GTTGGTGGAC 1500
CAAAATGAAG TCTACAGCAG GGCCTTCTAC TCCTGCGCAG CTGGCATGGC CTGGCTGCT 1560
GTGTGCTCG CCTGTGTGAG ACCGTGTAAG ATGGGACTGT GCCAGCATCA TCACCTCAGT 1620
GAAACAAAGG TAGTGAGCCA TCGTGGGAG ACITTCAGG ACATACCTGA AGACTTCTG 1680
GAAATGGAT TTGCAAAAA TGAGCACAGA GTTCACGTGC AAATGGAGCC GGTATGACAC 1740
ACTTCTTATC AACCAACAGC ACTGTGTTGG CTGGAGAGGG ATGGGGTGGG CCCAACGGG 1800
ACACAAGGAG GCAGAGGAGC TAACCCCTCT ACTCCACTTT CAAAACCTACA TTTTAAAGGG 1860

```

AATGTGTATG TGAAGAGCAC TACCAACATC GCTTTTGTIT TGTITTAAGC 1920
 TTTTITTTTT TGTITGTTTT TAAAGCCAAA ACAAAAAACA ACCAAGCACT CTTCATATA 1980
 TAAATCTGGC TGTATTCTGT AGCAATACAA GAGATATGTA GAAAGACTCT TTGGTTCACA 2040
 TCCGATATT AAAATAGTGA CATGAACATG CAAAGTGGIT TTAAGAGCTT TCACGTGGGA 2100
 TAAATGATT TCTTTTTC TTTTCTTTCT TCCTATGGTC TTGTCGAAAT AAACACTCT 2160
 CCTGAATAAA ACAACATCCA ACCCAGGTCA TTGAAATGAA ATTGGCCAGT C

Seq ID NO: 70 Protein sequence
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 | | | | |
 MTQNKLLKCS KANVYTEVPD GGWGNAAVAVS FFFVEVETYG LIKTGUVFFN DLMDSFNESN 60
 SRISWIIISIC VFVLTFSAPL ATVLSENRFQH RLVMVLGGLL VSTGMVAASF SQEVSHMYVA 120
 IGIISGLGYC FSLFTVTIL SQYFGKRRSI VTAVASTGEC FAVFAPAPAI MALKERIGWR 180
 YSLLFVGLLQ LNVIFGALL RPIIRGPAS PKIVIQENRK EAQYMLENEK TRTSIDSIDS 240
 GVELTTSFKN VPTHMLELE PKADMQQVLV KTSERPSEKK APLLDPSILK EKSFIYALF 300
 GLPATLGFFA PSLYIIPLGI SLGIDQDRAA FLLSTMAIAE VFGRIAGGFV LNRPIRKIY 360
 IELICVILLT VSLFAFTFAT EFWGLMCSOI FFGFMVGTIG GTHIPLLAED DVVGIEKMSS 420
 AAGVYIFIQS IAGLAGPPLA GLLVDQSKIY SRAFYSCAAG MALAAVCLAL VRPCKMGLCQ 480
 HHSGETKVV SHRGKTLQDI PEDFLEMOLA KNEHRVHVQM EPV

Seq ID NO: 71 DNA sequence
 Nucleic Acid Accession #: NM_004694
 Coding sequence: 166..1737

1 11 21 31 41 51
 | | | | |
 TTGGGGGTTT ATTCTCTTCC CTCTAACTT GACAGGGTCT TGCTCTGTCA TTCAGGCAAG 60
 AGTGCAGTAG TGTGATCACT TCTTACTGCC GCCTCAAGCT TCCAGCCTCA ACTCAAGCAA 120
 TCCCTCCACC TCAGCCACCC AAGTGGCTGG GACTACAGAT TAAGATGAC CCAAAATAAA 180
 TTAAGAGCTT GTTCCAAAGC CAATGTGTAT ACTGAAGTGC CTGATGGAGG ATGGGGCTGG 240
 GCGGTAGCTG TTTTATTTTT CTTCGTGAA GTCTTCACT ACAGCATCAT CAAGACATTT 300
 GGTGTCTTCT TTAATGACTT AATGGACAGT TTTAATGAAT CCAATAGCAG GATCTCATGG 360
 ATAATCTCAA TGTGTGTGTT TGTCTTAACA TTTTCAGCTC CCTCGCCAC AGTCTGTAGC 420
 AATGTCTTCG GACACCGTCT GGTAGTGTAT TTGGGGGGGC TACTTGTCTG CACCGGGATG 480
 GTGGCCGGCT CCTTCTCACA AGAGGTTTCT CATATGTACG TCGCCATCGG CATCATCTCT 540
 GGTCTGGGAT ACTGCTTTAG TTTTCTCCA ACTGTAAACA TCCTATCACA ATATTTTGGC 600
 AAAAGACGTT CCAATGTCAC TGCAGTTGCT TCCACAGGAG AATGTTTGGC TGTGTTTGGT 660
 TTCCACACAG CAATCTGTGC TCTGAAGGAG CGCATTTGGT GGAGATACAG CCTCTCTCTC 720
 GTGGCCCTAC TACAGTAAAA CATTGTCTAT TCCGAGGAC TGCTCAGACC CATCATATC 780
 AGAGGACGAG CGTCCACGAA AATAGTCATC CAGGAAATAC GGAAGAAGC GCAGTATATG 840
 CTGAAAAATG AGAAAACACG AACCTCAATA GACTCCATTG ACTCAGGAGT AGAAGTAACT 900
 ACCTCACCTA AAAATGTGCC TACTCACACT AACCTGGAAC TGGAGCCGAA GGCAGACATG 960
 CAGCAGGTCG TGTGTGAAGC CAGCCCCAGG CCAAGCGAAA AGAAAGCCCC GCTATTAGAC 1020
 TCTCTCATTT TGAAGAGAAA AAGTTTTATT TGTATATCAT TATTTGCTCT CTTTGCAACA 1080
 CTGGATCTCT TGCACCTTTC CTGTATATC ATTCTCTGG GCATTAGTCT GGGCATTGAC 1140
 CAGGACCGCG CTGCTTTTTT ATTATCTACG ATGGCCATTG CAGAAGTTTT CGGAAGGATC 1200
 GGAGCTGGTT TGTCTCTCAA CAGGGAGCCC ATTCTAAGA TTACATTTGA GCTCATCTGC 1260
 GTCATCTTAT TGAATGATTC TGTGTTTGGC TTTACTTTTG CTACTGAATT CTGGGGTCTA 1320
 ATGTATGCA GCATATTTTT TGGGTTTATG GTTGAACAA TAGGAGGACT CACATTCAC 1380
 TCTGTCTGCA AGATGATGTC GTGGCATG CAGAAGATGT CTCTGTCAGC TGGGGTCTAC 1440
 ATCTTCATTC AGAGCATAGC AGGACTGGCT GGACCGCCCC TTGCAGGTTT GTTGGTGGAC 1500
 CAAAGTAAGA TCTACAGCAG GGCTCTTCTC TCTGTGGCAG CTGGCATGGC CTGGCTGCT 1560
 GTGTGCTGCG CCTGGGTGAG ACCGTGTAAG ATGGGACTGT GCCAGCTCA TCACTCAGT 1620
 GAAACAAAGG TAGTGAGCCA TCGTGGGAG ACTTTACAGG ACATACCTGA AGACTTTCTG 1680
 GAAATGGATC TTGCAAAAAA TGAGCACAGA GTTCACTGTC AATGGAGGCC GGTATGACAC 1740
 ACTTTCTTAC AACCAACGCC ACTGTGTTGG CTGGAGAGGG ATGGGGTGGG CCCAACGGGG 1800
 ACACAGGAG GCACAGGAGC TAACCCCTCT ACTCCACTTT CAAAACATCA TTTTAAAGGG 1860
 AATGTGTATG TGAAGAGCAC TAACCAATC GCTTTTGTIT TGTITTTGTT TGTITTAAGC 1920
 TTTTITTTTT TGTGTTTTT TAAAGCCAAA ACAAAAAACA ACCAAGCACT CTTCATATA 1980
 TAAATCTGGC TGTATTCTGT AGCAATACAA GAGATATGTA GAAAGACTCT TTGGTTCACA 2040
 TCCGATATT AAAATAGTGA CATGAACATG CAAAGTGGIT TTAAGAGCTT TCACGTGGGA 2100
 TAAATGATT TCTTTTTC TTTTCTTTCT TCCTATGGTC TTGTCGAAAT AAACACTCT 2160
 CCTGAATAAA ACAACATCCA ACCCAGGTCA TTGAAATGAA ATTGGCCAGT C

Seq ID NO: 72 Protein sequence
 Protein Accession #: NP_004685

1 11 21 31 41 51
 | | | | |
 MTQNKLLKCS KANVYTEVPD GGWGNAAVAVS FFFVEVETYG LIKTGUVFFN DLMDSFNESN 60
 SRISWIIISIC VFVLTFSAPL ATVLSENRFQH RLVMVLGGLL VSTGMVAASF SQEVSEMYVA 120
 IGIISGLGYC FSLFTVTIL SQYFGKRRSI VTAVASTGEC FAVFAPAPAI MALKERIGWR 180
 YSLLFVGLLQ LNVIFGALL RPIIRGPAS PKIVIQENRK EAQYMLENEK TRTSIDSIDS 240
 GVELTTSFKN VPTHMLELE PKADMQQVLV KTSERPSEKK APLLDPSILK EKSFIYALF 300
 GLPATLGFFA PSLYIIPLGI SLGIDQDRAA FLLSTMAIAE VFGRIAGGFV LNRPIRKIY 360
 IELICVILLT VSLFAFTFAT EFWGLMCSOI FFGFMVGTIG GLTFKLLKM MSWALQKMS 420
 AAGVYIFIQS IAGLAGPPLA GLLVDQSKIY SRAFYSCAAG MALAAVCLAL VRPCKMGLCQ 480
 HHSGETKVV SHRGKTLQDI PEDFLEMOLA KNEHRVHVQM EPV

Seq ID NO: 73 DNA sequence
 Nucleic Acid Accession #: NM_002184.1
 Coding sequence: 256..3012

1 11 21 31 41 51
 5 GAGCAGCCAA AAGGCCCCCG GAGTCGCGCT GGGCCGCCCC GGCGCAGCTG AACCGGGGGC 60
 CGCGCCTGCC AGGCCGACGG GTCTGGCCCA GCCTGGCGCC AAGGGGTTCT TGCGCTGTGG 120
 AGACGCGGAG GGTGCGAGCG GCGCGGCCCT AGTGAACCC AATGGAAAA GCATGACATT 180
 TAGAAGTAGA AGACTTAGCT TCAATCCCT ACTCCTTCAC TTACTAATT TGTGATTGG 240
 AAATATCGC GCAAGATGTT GACCTTGCAG ACTTGGGTAG TGCAGCCCT GTTTATTTTC 300
 CTCACCACTG AATCTACAGG TGAACCTCTA GATCCATGTG GTTATATCAG TCCTGAATCT 360
 CCAGTTGTAC AACTTCATT TAATTTCACT GCAGTTTGTG TGTAAAGGA AAAATGTATG 420
 10 GATTATTTTC ATCTAAATGC TAATTACATT GTCTGGAATA CAAACCAATT TACTATTCTT 480
 AAGGAGCAAT ATACTATCAT AAACAGAACA GCATCCAGTG TCACCTTTAC AGATATAGCT 540
 TCATTAATA TTGAGCTCAC TTGCAACATT CTACATTGG GACAGCTTGA ACAGATGTT 600
 TATGGAATCA CAATAATTTC AGGCTTGCCT CCAGAAAAAC CTAAAAATT GAGTTGCATT 660
 GTGAACGAG GGAAGAAAAAT GGGGTGTGAG TGGGATGTTG GAAGGGAAAC ACACPTGGAG 720
 15 ACAAACTTCA CTTTAAATAT TGAATGGGCA ACACACAGT TTGCTGATTG CAAAGCAAAA 780
 CGTGACACCC CCACCTCATG CACTGTTGAT TATTCTACTG TGTATTTTGT CAACATTGAA 840
 GTCTGGGTAG AAGCAGAGAA TGCCTTTGGG AAGGTTACAT CAGATCATAT CAATTTTGT 900
 CCTGTATATA AAGTGAAGCC CAATCCGCCA CATAATTAT CAGTATCAA CTCAGGGGAA 960
 20 CTGCTAGTAT TCTTAAATTT GACATGGACC AACCACAGTA TTAAGAGTGT TATAATACTA 1020
 AAATATAACA TCTAATATAG GACCAAGAT GCCTCAACTT GGAGCCAGAT TCCTCTGTAA 1080
 GACACAGCAT CCACCCGATC TTCACTCACT GTCCAGAGCC TTAACCTTTT TACAGAATAT 1140
 GTGTTTAGGA TTGCTGTAT GAAGGAAGAT GGTAAAGGAT ACTGGAGTGA CTGGAGTGAA 1200
 GAAGCAAGTG GATCACCTA TGAAGATAGA CCATCTAAAG CACCAAGTTT CTGGTATAAA 1260
 25 ATAGATCCAT CCCATACCTA AGGCTACAGA ACTGTACAACT TCGTGTGGAA GACATTGCTT 1320
 CCTTTTGAAG CTTTAAATAT AATCTTGGAT TATGAAGTGA CTCTCACAAG ATGGAATCAA 1380
 CATTACAAAT ATACACAGT TAATGCCACA AAACGTACAG TAAATCTCAC AAATGATCGC 1440
 TATCTAGCAA CCTTACAGT AAGAAATCTT GTTGGCAAT CAGATGCAGC TGTTTTAACT 1500
 ATCCCTGCCCT GTGACTTTCA AGCTACTCAC CCTGTAATGG ATCTTAAAGC ATTCCCAAAA 1560
 30 GATAACATGC TTTGGGTGGA ATGGACTACT CCAAGGGAAT CTGTAAAGAA ATATATACTT 1620
 GAGTGGTGTG TGTATGAGA TAAAGCACCC TGTATCACAG ACTGGCAACA AGAAGATGGT 1680
 ACCGTGCATC GCACCTATTT AAGAGGGAAC TTAGCAGAGA GCAATGCTA TTTGATAACA 1740
 GTTACTCCAG TATATGCTG TGGACAGGA AGCCCTGAAT CCNTAAAGGC ATACCTTAAA 1800
 CAAGCTCCAC CTTCCAAAGG ACCTACTGTT CGGACAAAA AAGTAGGGAA AAACGAAGCT 1860
 35 GTCTTAGAGT GGGACCAACT TCCTGTTGAT GTTCAGATG GATTATCAG AAATTATACT 1920
 ATATTTTATA GAAACCATAT TGGAAATGAA ACTGCTGTGA ATGTGGATT TCCTCCACAA 1980
 GATATATACAT TGTCTCTTT GACTAGTGAC ACATTGTACA TGGTACGAAT GGCAGCATAC 2040
 ACAGATGAAG GTGGGAAGGA TGGTCCAGAA TTCACTTTTA CTACCCCAA GTTGTCTCAA 2100
 GGAGAAATG AAGCCATAGT CGTGCCCTGT TGTCTAGCAT TCCTATTGAC AACTCTTCTG 2160
 40 GGAGTGCTGT TCTGCTTTAA TAAGCGGAGC CTAATTAAAA AACACATCTG GCCTAATGTT 2220
 CCAGATCCCT CCAAGCATAT TATTGCCAG TGGTCACTTC ACACCTCTCC AAGGCACAT 2280
 TTTAATTCAA AAGATCAAT GTATTCCAGT GGCAATTTCA CTGATGTAAG TGTGTGGAA 2340
 ATAGAAGCAA AAGTCAAAA GCCTTTTCCA GAAGATCTGA AATCATTTGA CCTGTTCAA 2400
 AAGGAAAAAA TTAATACGTA AGGACACAGC AGTGGTATTG GGGGCTCTTC ATGCATGTCA 2460
 45 TCTTCTAGC CAAGCATTTT TAGCAGTGAT GAAATATGAT CTTCACAAA CACTTCGAGC 2520
 ACTGTCCAGT ATCTTACCGT GGTACACAGT GGTACACAGC ACCAAGTTCC GTCACTCAA 2580
 GTCTTCTCAA GATCCGAGTC TACCACGCC TTGTTAGATT CAGAGGAGCG GCCAGAGAT 2640
 CTACAATTAT TATATCATG AGATGGCGGT GATGGTATT TGGCCAGGCA ACAGTACTTC 2700
 AAACGAACCT GCAGTCAGCA TGAATCCAGT CCAGATATTT CACATTTTGA AAGGTCAAAG 2760
 50 CAAGTTTCAT CAGTCAATGA GGAAGATTTT GTTAGACTTA AACAGCAGAT TTCAGATCAT 2820
 AITTCACAT CCTGTGGATC TGGGCAAAAT AAAATGTTTC AGGAAGTTTC TGCAGCAGAT 2880
 GCTTTTGGTC CAGGTACTGA GGGCAAGTA GAAAGATTG AAACAGTTGG CATGGAGGCT 2940
 GCGACTGATG AAGGCATGCC TAAAGTTAC TTACCACAGA CTGTACGGCA AGGCGGCTAC 3000
 55 ATGCTTCAGT GAAGGACTAG TAGTCTCTGC TACAACCTCA GCAGTACCTA TAAAGTAAAG 3060
 CTAATATGAT TTTATCTGTG AATTC

Seq ID NO: 74 Protein sequence
 Protein Accession #: NP_002175.1

60 1 11 21 31 41 51
 MLTLQTVVVQ ALFIFLTIES TGLLDPCGY ISPESPVVQL HSNFTAVCVL MKKMDYFHV 60
 NANYIVKKTN HPTIPKEQYT IINRTASSVT FTDIASLNIQ LTCNLITPGQ LEQNVYGITI 120
 65 IGLPPEKPK NLSCIVNBSK RMRCEWDGGR ETHLEINFTL KSEWATHKFA DCKAKRDTFT 180
 SCTVDYSTVY FVNIEVWVEA ENALGKVTSD HINFPVYKV KPMFPNLSV INSEELSSIL 240
 KLTWNPISIK SVIILKYNIQ YRTKDASTNS QIPPEDTAST RSSFTVQDLK PFTYVFRIR 300
 CMKEDGKGXW SDWSEASGI TYEDRPSKAP SFWYKIDPSH TQGYRTVQLV WKTLPPFEAN 360
 GKILDYEVTI TRWKSHLQNY TVNATKLTVN LTNDRYLATL TVRNLVGKSD AAVLTIRACD 420
 70 PQATHFVMDL KAFPKDMLM VEWTFPRSSV KKYILENCVL SDRAPCITDW QQEDGTVBRT 480
 YLRGNLAESK CYLITVTPVY ADGPGSPESI KAYLEQAPPV NGFTVRIKKV GKNEAVLEWD 540
 QLPVDVQNGP IRNYTIFYRT IIGNETAVNV DSSETEYTL SLSDTLYMV RMAAYTDEGG 600
 KDGPSPFTT PKFAQGLEIA IVVPVCLAFI LTTLLGVLFK FNKRDLIKH IWPNVDPFSK 660
 SHLAQWSPHT PRRHNFNSKD QMSYDGNFTD VSVVEIAND KKFPPDLKS LDLFKKEKIN 720
 75 TEGHSSGIGG SSCMSSSRPS IBSSENEBS QNTSSTVQYS TVVHSGYRHO VPSVQVPSRS 780
 ESTQPLDSE ERPEDLQLVD HVDGGDGLIF RQYFKQNCB QHSSPDISH PERSKQVSSV 840
 NEEDFVRLKQ QIDSHISQSC GSGQMKMPQB VSAADAFPGP TEGQVERFET VGMERATDEG 900
 MPKSYLPQTV RQGGYMEQ

Seq ID NO: 75 DNA sequence
 Nucleic Acid Accession #: NM_022131
 Coding sequence: 11..2878

80 1 11 21 31 41 51
 TGTGCGAGG ATGCTGCCTG GCGCGCTGTG CTGGGTGCGG CTCTGCTGCG CCGTGGCGGT 60

5	GGGGAGCGGC	AGCGGCGGTG	GCGGGGACAG	CGGGCAGGCG	CGCTCCTCG	CGGCTAAAGT	120
	CAATAAGCAC	AAGCCATGGA	TCAGAGCTTC	ATATCATGGA	GTCTAAGCTG	AGAACAATGA	180
	CACAGTCATT	TTGGACCCAC	CACCTGGTAGC	CCTGGATAAA	GATGCACCGG	TTCCCTTTTGC	240
	AGGGGAAATC	TGTGCGTTC	AGATCCATGG	CCAGGAGCTG	CCCTTTGAGG	CTGTGGTGCT	300
	CAACAAGACA	TCAGGAGAGG	GCGGCTCCG	TGCCAAGAGC	CCCATTTGACT	GTGAGTTGCA	360
	GAAGGAGTAC	ACATTCATCA	TCAGGCTTA	TGACTGTGGT	GCTGGGCCCC	ACGAGACAGC	420
	CTGGAAAAAG	TCACACAAGG	CCGTGGTCCA	TATACAGGTG	AAGGATGTCA	ACGAGTTTGC	480
	TCCACCTTC	AAAGAGCCAG	CCTACAAGGC	TGTTGTGACG	GAGGGCAAGA	TCATGACAG	540
10	CATTCTGACG	GTGGAGGCCA	TTGACAGGGA	CTGCTCCCA	CAGTACAGCC	AGATCTGCAA	600
	CTATGAATAT	GTCCACACAG	ATTGCTCTT	TGCCATCGAC	AGAAATGGCA	ACATCAGGAA	660
	CACTGAGAA	CTGAGCTATG	ACAAACAACA	CCAGTATGAG	ATCCTGGTGA	CCGCTTAAGA	720
	CTGTGGACAG	AAGCCCGCTG	CTCAGGACAC	CCTGGTGCAG	GTGGATGTGA	AGCCAGTTTG	780
	CAAGCCTGGC	TGGCAAGACT	GGACCAAGAG	GATTGAGTAC	CAGCCTGGCT	CCGGGAGCAT	840
15	GCCCCCTGTC	CCCGAGCTCC	ACCTGGAGAC	GTGGATGGA	GCCGTGTCTT	CCCTCCAGAT	900
	CGTCACAGAG	CTGCAAGCTA	ATTACATTGG	GAAGGTTGT	GACCGGAGA	CTTACTCTGA	960
	GAATATCCCT	CAGAGTTTAT	GTGGAGCCTC	CTCTGGCTTC	ATTGAGCTCT	TGCCATCCCC	1020
	TAGCGCTGCC	ACCACTGGA	CTGCAGGACT	GCTGGTGGAC	AGCAGTGA	TGATCTTCAA	1080
	GTGTGACGCG	AGGACAGGTG	CCAAATCC	CGATGGGATT	GTGCCCAAGA	ACCTGACCGA	1140
20	TCAGTTCAC	ATCACCATGT	GGATGAACA	CGGCCCCAGC	CCTGGTGTGA	GAGCCGAGAA	1200
	GGAAACCATC	CTCTCCACT	CAGACAAAC	CGAAATGAAC	CGGCATCACT	ATGCCCTGTA	1260
	TGTGCACAA	TGCCGCTCC	TCTTCTCTT	CGGGAAGGAC	TTCGACCCAG	CTGACACCTT	1320
	TGCCCCCGCG	GAGTTCCTCT	GGAGGCTGGA	TCAGATTGT	GACAAAGAGT	GGCACTACTA	1380
	TGTATCAAT	TGGGAGTTTC	CTGTGGTAAC	CTTATACATG	GATGGAGCAA	CATATGAACC	1440
25	ATACCTGGG	ACCAACGACT	GGCCCATTTA	TCCATCTCAC	ATAGCCATGC	AACTCAGAGT	1500
	CGCGCTGCT	TGGCAAGGAG	GAGAAGTCAC	CAAAACACAG	TTTGTCTAGT	TCCTTCATGG	1560
	AAGCCCTGGC	AGTCTCACCA	TCCGCCCTGG	CAAAATGGA	AGCCAGAGG	TGATCTCCCT	1620
	CCTGCAAGGC	TGCAAGGAG	GGCTGGACAT	TAATTCCTTG	GAAAGCCTTG	GCCAGGAAAT	1680
	AAAGTATCAC	TTCAACCCCT	CGCAGTCCAT	CCTGGTGATG	GAAGGTGACG	ACATTTGGAA	1740
30	CATTAACCTG	GCTCTCCAGA	AAGTCTCCTA	CATCAACTCC	AGGCAGTTCC	CAACGCGCGG	1800
	TGTGCGGCG	CTCAAGGAG	CCCTCCAAAGT	CCAGTCTTT	GGGGAAGACG	TATGCATCAG	1860
	TATCCCTGAG	GTAGATGCTT	ATGTGATGGT	CCTCCAGGCC	ATCGAGCCCC	GGATCACCCT	1920
	CGGGGACACA	GACCACTTCT	GGAGACCTGC	TGCCAGTTT	GAAAGTGCCA	GGGGAGTGAC	1980
	CCTCTCCCT	GATATCAAGA	TGTGAGCAC	CTTCCCAAA	ACCGAAGCCC	CGGGGACGCT	2040
35	GAAACACCA	GACCCCAAAT	CAGAACTCT	AGAGGAAATG	CTTCATAACT	TAGATTCTTG	2100
	TGACATTGTT	GTGATGGGAG	GGGACTTGGA	CCCAAGGCG	GAGTGTCTGG	AGCTCAACCA	2160
	CAGTGAGCTC	CACCAACGAC	ACCTGGATGC	CATTAATCT	ACTGCAGGCT	ACTCCATCTA	2220
	CGGTGTGGGC	TGCTATGAGC	GCTATGAGCA	GGTGTACAT	CACATCCGCT	ACCGCAACTG	2280
40	GCGTCCGCT	TCCCTTGAGG	CCCGGCTTT	CGGATTAG	TGCTCAGAAC	TCAATGGGCG	2340
	CTACACTAGC	AATGAGTTCA	ACTTGGAGGT	CAGCATCTT	CATGAAGACC	AAGTCTCAGA	2400
	TAGAGCTGAG	GTCAATCATC	TGATTGTGCA	GCTTCCCTTC	CTCCAGTCTG	TCCATCATCC	2460
	TGAGTCCCG	AGTAGCTCC	AGCAGAGTTC	AGTGTCCCA	AGCATTGCCA	CAGTGGTCAT	2520
	CATCATCTCC	GTGTCTATGC	TGTGTGTTGT	CGTGGCCATG	GGTGTGTACC	GGGTCCGGAT	2580
	CGCCACACAG	TGCTGCTATC	AGGAGACTGA	GGCTGCCAAG	GATCTGAGA	TGGACTGGGA	2640
45	CGATTCTGG	CTGACTATCA	CAGTCAACCC	CATGGAGAAA	CATGAAGGAC	CAGGGCATGG	2700
	GGAAGATGAG	ACTGAAGGAG	AAGAGGAGGA	AGAAGCCGAG	GAAGAAATGA	GCTCCAGCAG	2760
	TGGCTCTGAC	GACAGGGAAG	AGGAGGAGGA	GGAGGAAGGG	ATGGGCAGAG	GCAGCATGAG	2820
	GCAGAAATGA	GCCAGGCAAG	CCCAGCTGGA	GTGGGATGAC	TCCACCTTCC	CCTACTAGTG	2880
	CCCAAGGCTC	TGCTGCTGG	CCCACTGTC	CCCTTTGTAA	ACCCTGACCC	AGTGTATGCC	2940
50	CATGCTATC	ATACCTTACC	TCTGATGCT	GTGACATGTC	TGGGAAGGCC	TCTCCAGCT	3000
	TCCGTGGGCC	CACCTTTTAA	GCTTGGGCA	CTCCCTGTGT	TTCTATCCATG	GGGAAGTTCC	3060
	AAGAAGCCCA	GCTATGGCAT	CAGTGAGGAC	TTCAGGGTAG	ACTTGTCTCT	GTAGCCTCCA	3120
	CTTCTGCTCT	AAGTTCCCCA	GCATCTGAC	TACCTGCTG	CAGAGTTTGC	CTTTGTTTT	3180
	TCTTGCAGGG	AAGAAGGCC	ACCTTTGTGT	CATCACTCTC	CCCAGGCTCA	GAGTCCCAAA	3240
55	GGCCCTGGG	TTCCAACTCA	CTGTGCGTCT	CCTCCACACA	GACCACTAGG	TCTCTCTATG	3300
	CTGACTCCAG	TTGTCTCTAT	ACAAGGAGGG	TGGTTGAACT	TCACACACGT	AAGGTCTTAG	3360
	TGCTTAACAG	TTTAAAGGAA	AGTCTTGT	CAGGCAGAAC	TAGTTTACA	GGGAAAGGTA	3420
	CACACATCT	CTCTCTCTCT	CTCTCTCTGT	CTATCTAGTT	CCCAAGCTTG	GAGAGCCTTT	3480
	CCCTTCTGCT	CTTCTGAGG	CCATATAAGC	TTATAAGAA	AGTCCCAAC	CAAGAATAGG	3540
60	TCCCTGGCCA	CAGCAGGGT	CTGATCCCTC	ATCAGAGCTA	TCTGAGCCTG	CCTGTCTGGG	3600
	CACCTGTGTC	AACCATGCG	CTACCTGCTC	AGGGGCACTC	AGCAACAGA	ACCCAGGGGC	3660
	CCAGGAGGCA	TTCACACAG	GCACTGCCCT	AGGACACAC	AACAGGACA	GTCAACACAA	3720
	GGACAAACAG	GACCAACAC	AACACACAC	AGGACAGTC	ACAACAAGCC	TAGAGCCAGA	3780
	AAGCAGATGG	AAATGCTAAT	GAGGTCAAAC	GTAGGCTTCA	TGGTGGGTGG	AGTGGGGGTG	3840
65	GCTGGGCTCC	CCCAGGACAG	AGGGGACCTT	GAGGTGGGCA	AGGCTCTCAC	CACCTCAGCT	3900
	TATGGTCCCT	CTTCTCTCT	CTTCCCTCT	GAGAAATAC	ACGCTTCTG	CATGTATTAG	3960
	AAACGCAAGA	GCTCCACCAA	GTCTACAATG	AAAGTTTGA	ATTAACTG	AAGGAATTAG	4020
	AGCATATTT	GCAATCATTG	CAGCTTCTTC	TTTCTTCTG	TCATAAAGG	AGGAACACTT	4080
	TAGATAGAGG	GCAATATAT	CTGAACACT	AATTTCTTC	TTTTTTTAT	AAGGAATCT	4140
70	TTTCCATCTC	CATCTAACA	TGCACACCT	GTGAAGAGAA	TTGTTTCTAT	AGTAATCTGT	4200
	CTGTGATCTT	TTGTGGCCAA	GAGAATAGCA	GGCAAGAAAT	AGGGCCCTGA	CAGAATTTCC	4260
	ACGAAGCTCT	GAGAACATGT	TGTTTTCGAA	TGTCTGATTC	CTCTTTGTCA	TCAATGTGTA	4320
	TGCTCTGTCC	CCATCTCTCA	CTCCTCTCTA	AGCTCACACC	AATTGGTTTG	GCACAGGCAC	4380
	AGAGCTGGTC	CCTAGTTAAG	TGGCATTTAT	GTAAAAAA	A		
75	Seq ID NO: 76 Protein sequence						
	Protein Accession #: NP_071414						
80	1	11	21	31	41	51	
	MLPGRLCNVP	LLALGVGSG	SGGGGDSRQR	RLLAAKVNKH	KPWIETSYRG	VITENNNTVI	60
	LDPELVALDK	DAPVPFAGBI	CAFKINGQEL	PFEAVVINKT	SGEGRRLAKS	PIDCKLQKEY	120
	TFIIQAYDCG	AGPHEIANKK	SHKAVVHIQV	KDVNEFAPTF	KEPAYKAVVT	EGRIYDSILQ	180
	VFAIDEDCBF	QYSGICNTXEI	VTDVFPFALD	RNGNLIRTEK	LSYDRQHQYE	ILVTAYDCGQ	240
	KPAAQDTLVQ	VDVKPKVCKEG	WQDNTKRLFY	QPGSGSMPLF	PSIHLETCDG	AVSBLQIVTE	300

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75

LQTNVIGKGC DRETYSEKSL QKLOGASSGI IDLLPSPSAA TNWTAGLLVD SSEMIFKFDG 360
 RQGAIPDGI VPJMLTDQFT ITNMKKGPS PGVRAEKETI LCNSDKTEMN RHHYALYVHN 420
 CRLVPLLRKD EDQADTFRPA EFHWKLDQIC DKWHYVYVIN VEPVVTLYM DGATYEPVLV 480
 TNDWPIHPSH IAMQLTVGAC WQGGEVTKPQ PAQPFHGSLSA SLTIRPGKME SQKVISCLQA 540
 CKEGLDINSL ESLQGGIKYH FNPSQILVM EDDIGNINR ALQKVSYSIN RQFPTAGVRR 600
 LKVSSEKVOCF GEDVCISPE VDAYVMVLQA IEPRIITLRGT DHFWRPAAQF ESARGVTLFP 660
 DIKIVSTFAK TEAPGDVKT DPKEVLEEM LHNLDFTDIL VIGGDLOPRQ ECLELNESEL 720
 RQRHLDATNS TAGYSIYGVG SMSRYEQVLH HIRYRNWRPA SLEARRFRK CSZLNGRYTS 780
 NEFNLEVSIH HEDQVSDKEE VNHLTVQPPF LQSVHHPESR SSIQHSVVVP SIATVVIIS 840
 VCLVFFVAM GVYFVRIBAQ HFIQETEAAK ESEMOWDDSA LTITVNMEEK HEGPGHGEDE 900
 TGESEEEAE REMSSSSQSD DSEEEEEES MGRGRHGQNG ARQAQLEWDD STLPY

Seq ID NO: 77 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 482..3007

1 11 21 31 41 51
 AACTGAGCTA ACAAGAAATA CTAGAAAAGG AGGAAGGAGA ACATTGCTGC AGCTTGGATC 60
 TACAACCTAA GAAAGCAAGA GTGATCAATC TCAGCTCTGT TAAACATCTT GTTTACTTAC 120
 TGCAATTCAGC AGCTTGCAAA TGGTTAACTA TATGCAAAAA AGTCAGCATA GCTGTGAAGT 180
 ATGCCGTGAA TTTTAAATTGA GGGAAAAGG GACAATTGCT TCAGGATGCT CTAGTATGCA 240
 CTCTGCTTGA AATATTTTCA ATGAAAATGCT CAGTATTCTA TCTTTGACCA GAGGTTTAA 300
 CTTTATGAAG CTATGGGACT TGACAAAAGG TGATATTGA GAAGAAAGTA CGCAGTGGT 360
 GGTGTTTCT TTTTAAAT AAAGGAATG AATTACTTTG AACACCTCTT CCAGCTGTGC 420
 ATTACAGATA ACGTCAGGAA GAGTCTCTGC TTTACAGAAT CGGATTTTCA CACATGACAA 480
 CATGAAGCTG TGGATTCTAT TCTTTTATC ATCTCTCTT GCCGTATAT CTTTCACTC 540
 CCAACTCCA GTGCTCTCAT CCAGAGGCTC TTGTGATTCT CTTTGCAATT GTGAGGAAA 600
 AGATGGCACA ATGCTAATAA ATGTGGAAGC AAAAGGTATC AAGATGGTAT CTGAAATAAG 660
 TGTGCCACCA TCCAGACTT TCCAACTAAG CTTATTAAAT AACGGCTTGA CGATGCTTCA 720
 CACAAATGAC TTTTCIGGGC TTACCAATGC TATTTCAATA CACCTTGGAT TTAACATAT 780
 TGCAGATATT GAGATAGGTG CATTAAATG CTTGGCCTC CTGAAACAAC TTTATATCAA 840
 TCACAAATCT TTAGAAATTC TAAAGAGGA TACTTTCAT GGACTGGAAA ACCTGGAATT 900
 CCTGCAAGCA GATAACAATT TTATCAGAGT GATTGAACCA AGTGCCCTTA GCAGGCTCAA 960
 CAGACTCAAA GTTTAATTT TAAATGACAA TGTATTGAG AGTCTTCTCT CAACATCTT 1020
 CCGATTGTT CTTTAAACCC ATCTAGATCT TCTGTGAAAT CAATTACAAA CATGCGCTTA 1080
 TGTGTGTTT CTCGAACACA TTGGCCGAAT ATTGGATCTT CAGTTGGAGG ACACAAATG 1140
 GGCTGCAAT TGTGACTTAT TGCAGTTAAA AACTTGGTTG GAGAACATGC CTCACAGTC 1200
 TATAATTGGT GATGTTGCT GCAACAGGCC TCCATTTTTT AAAGGAAGTA TACTCAGTAG 1260
 ACTAAGAGAG GAATCTATTT GCCCTACTCC ACCAGTGTAT GAAGAACATG AGGATCCTTC 1320
 AGGATCATTA CATCTGGCAG CAACATCTTC AATAAATGAT AGTCGCATGT CAACTAAGAC 1380
 CAGGTCCATT CTCAAACTAC CCACCAAAGC ACCAGGTTTG ATACCTTATA TTCAAAAGCC 1440
 ATCCACTCAA CTTCCAGGAC CTTACTGCCC TATCTCTGT AACTGCAAG TCCATATCCC 1500
 ATCAGGACTT CTAATACATT GTCAGGAGCG CAACATTGAA AGCTTATCAG ATCTGAGACC 1560
 TCCCTCCGAA AATCCCTAGAA AGCTCATCT AGCGGGAAAT ATTATTCA CA GTTTAATGAA 1620
 GTCGATCTA GTGGAATATT TCACCTTGGG AATGCTTAC TTGGGAAACA ATCTATTTGA 1680
 AGTCTCTGAA GAAGATCGT TTATGAACCT AACGAGATTA CAAAACCTCT ATCTAAATGG 1740
 TAACCACCTG ACCAAATTAA GTAAAGGCAT GTTCTTTGGT CTCCATAATC TTGAATACTT 1800
 ATATCTTGA TACAATGCCA TTAAGGAAAT ACTGCCAGGA ACCTTAATC CAATGCTTAA 1860
 ACTTAAAGT CTTTATTATA AATAACAACCT CTTCCAGTT TTACCAACAC ATATTTTTTC 1920
 AGGGGTTCTT CTAACTAAG TAAATCTTAA AACAAACCGG TTTACCCATC TACCTGTAAG 1980
 TAAATATTGG GATGATCTTG ATTACTAAC CCAGATTGAC CTTGAGGATA ACCCTGGGA 2040
 CTGCTCTCTG GACCTGTTG GACTGCAACA ATGGATACAA AAGTTAAGCA AGAACACAGT 2100
 GACGATGAC ATCTCTGCA CTTCCCCCGG GCATCTGAC AAAAGGAAT TBAAGCCCT 2160
 AAATAGTGA ATCTCTGTC CAGGTTTATG AATAAACCCA TCCATGCCAA CACAGACTAG 2220
 TTACCTTATG GTCACCACTC CTGCAACAAC AACAAATACG GCTGATACTA TTTTACGATC 2280
 TCTTAAGGAC GCTGTGCCAC TGTCTGTTCT AATATTGGGA CTTCTGATTA TGTTCATCAC 2340
 TATGTTTTC TGTGCTGCAG GATAGTGTGT TCTGTTCTT CACCGCAGGA GAGATACAA 2400
 AAAGAAACAA GTAGATGAGC AAATGAGAGA CAACAGTCTT GTGCACTTTC AGTACAGCAT 2460
 GTATGGCCAT AAAACCACTC ATCACACTAC TGAAAGACCC TCTGCCTCAC TCTATGAACA 2520
 GCACATGGTG AGCCCCATGG TTCACTGCTA TAGAAGTCCA TCTTTGGTTC CAAAGCATCT 2580
 GGAAGAGGAA GAAGAGAGGA ATGAGAAAGA AGGAAGTGAT GCAAAACATC TCCAAAGAAG 2640
 TCTTTTGGAA CAGGAAATC ATTCAACACT CACAGGGTCA AATATGAAAT ACAAACAC 2700
 GAACCAATCA ACAGAAATTT TATCCTTCCA AGATGCCAGC TCATTGTACA GAAACATTT 2760
 AGAAAAGAA AGGGAACCTC AGCAACTGGG AATCACAGAA TACCTAAGGA AAAACATTGC 2820
 TCAGCTCCAG CCTGATATGG AGGCACATTA TCTGGAGGCC CACGAAGAGC TGAAGTTAAT 2880
 GGAACATTA ATGTACTCAC GTCCAAGGAA GGTATTAGTG GAACAGACAA AAANTGAGTA 2940
 TTTTGAACCT AAAGCTAATT TACATGCTGA ACCTGACTAT TTAGAAGTCC TGGAGCAGCA 3000
 AACATAGATG GAGAGTTTGA GGGCTTTCGC AGAATGCTG TGAATCTGTT TTAAGTCCAT 3060
 ACCITGTAAA TAGTGCCCTT ACGTGAGTGT GTCATCAATC AGAACCTAAG CACAGCAGTA 3120
 AACTATGGGG AAAAAAAG AAGAGAAAA GAAACTCAGG GATCACTGGG AGAAGCCATG 3180
 GCATATCTT CAGCAATT AGTCTGCCC AAATAAATC AATCCITGCA TGTAAATC

Seq ID NO: 78 Protein sequence

Protein Accession #: Eos sequence

1 11 21 31 41 51
 MKLWHLFPYS SLIACISLHS QTPVLSSRGS CDSLNCREEK DGTMLINCSA KGIKMSVBSIS 60
 VPPSRPPQLS LLNGLTLMR TNDPFGSLTNA ISIHGFMNI ADIEIGAFNG LGLLQLHLIN 120
 HNSLEILKED TFGGLENLFP LQADNFIIV IEPFAPSKLN RLKVLILNDN AIESLPPNIF 180
 RVPPLTHLDL RGNQLQTLFY VGFLEHIGRI LDQLLEDNKA ACNCDLLQLK TLENMPPQS 240
 IIGDVVCSNP PFFKGSILSR LKBSICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT 300
 TSLILPLPRA PGLIPYITKP STOLPQPYCP IFCNCKVLSP SGLLIHQER NIESLSDLRP 360

PPQNPRKIL AGNIHSLMK SDIVEYFTLE MLHLGNRRIS VLREGSFMNL TRIQKLYLNG 420
 MHLTKL SKGM FLGLHNLBYL YLEYNALKEI LPQTFNPMFK LKVLILNNNL LQVLPPHIFS 480
 GVPLTKVNLK TNQFTHLFVS NILDLDLLT QIDLEDNPDW CSDLVGLQQ WIQKLSKNTV 540
 TDDILCTSPG HLDKKEKLAL NSEILCPGLV NNPSPMTQTS YLMVTTTATT TINTADTLRS 600
 LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KQVDEQMRD NSPVHLQYSM 660
 YGHKTTHHTT ERPSASLYEQ HMVSPMVHVY RSPSFGPKHL EEBBERNEKE GSDAKELQRS 720
 LLEQENHSPL TGSNMKYKTT NQSTPLSFQ DASSLYRNIL EKERELQQLG ITEYLKKNIA 780
 QIQPDMEAHY PGHSELEKLM ETIMYSRPRK VLVEQTKNEY FELKANLHAE PDYLEVLEQQ 840
 T

Seq ID NO: 79 DNA sequence

Nucleic Acid Accession #: NM_016640.2

Coding sequence: 39..1358

1 11 21 31 41 51
 | | | | |
 GCTTAAGTTG ACCCTCTGGT CCGGAATCGC GGGCAAAGAT GCGCGCGGCC AGGTGTGGGA 60
 GGCCTTTGCT ACGCGGTCCG AGGCTTTCAT TGCACACCGC GGCTAATGCC GCGCCACGG 120
 CTACAGAAAC GACCTCCCAA GACGTCGCGG CAGCCCGCGT CGCGCGGTAC CCGCCGATTG 180
 TGGCTCCAT GACAGCCGAC AGCAAAGCTG CACGGCTGCG GCGGATCBAG CGCTGGCAGG 240
 CGACGGTGCA CGCTGCGGAG TCGGTAGACG AGAAGCTGCG AATCTCTACC AAGATGCACT 300
 TTATGAAGTA CATGGTTTAC CCGCAGACCT TCGGCTGAA TCGCGACCGC TGGTACCAGT 360
 ACTTCACCAA GACCTGTGTC CTGTGCGGTC TCGCGCGGCC CCGAGCGGAG CCGGAGCCCG 420
 AGCCCGAACC CGAAGCTGAA CCTGCGCTGG ACCTCGCGGC GCTGCGTGGG GTGCGCTGGG 480
 ACTGCTGCT GACAGGAGCA TTCTACCTGC GCGCGAGGCG GCGCGTGAC CATTACGAGG 540
 AGAGCGAGGT CATATCTTTC CCTTCTCTGG ATCAGCTGGT GTCAACCTTC GTGGGCTCTC 600
 TCAGCCCA CAACCCGCGC CTGCGCGCTG CCGCCCTCGA TTATAGATGC CCAGTTCATT 660
 TTTACTGGGT GCGTGTGAAA GAATTTATTC CTGTTGGTCA TCGAAGAGGT CGAATTGATG 720
 ACTTGGGATA CCAGATAGAT GATAAACCAA ACAACAGAT TCGAATATCC AAGCAACTCG 780
 CAGAGTTTGT GCCATGGAT TATTCTGTTT CTATAGAAAT CCCCCTATA AAATGTAAAC 840
 CAGACAACT TCCATTATTC AAAGGCGAGT ATGAAACCA CATATTGTGT GGCTCAAAA 900
 CTGCGATCC TTGCTGTATC GGTCCACCCC AGTTTCATCT GTTACCTGAC AAATTAAGAA 960
 GGGAAAGGCT TTGTAGACAA AACTGTGCTG ATCAGATAGA AGTTGTTTT AGAGCTAATG 1020
 CTATTGCAAG CCTTTTGTCT TGGACTGGAG CACAAGCTAT GTATCAAGGA TTCTGGAGTG 1080
 AAGCAGATG TACTCGACCT TTTGTCTCCC AGGCTGTGAT CACAGATGGA AAATACITTT 1140
 CCTTTTCTG CTACAGCTA AATACTTTGG CACTGACTAC ACAAGCTGAT CAAATAACCC 1200
 CTCGTAATAA TATATGTTGG GGTACACAAA GTAAAGCTCT TTATGAACCA ATTGAGGATA 1260
 ATGATGTGAA AGGTTTAAAT GATGATGTTT TACTTCAGAT AGTTCACTTT CTACTGAATA 1320
 GACCAAAAGA AGAAAAATCA CAGCTGTGCG AAAACTGAAA AAGCATATTT GATTGAGAAC 1380
 TGTGGGAATA TTTAATTTT ACTGAAGGAA CAATAATGAT GAGATTGTA ACTGTCAACT 1440
 ATTAATATCA TTGATTTTGG AGACAAAAA AAAAAAAAAA AA

Seq ID NO: 80 Protein sequence

Protein Accession #: NP_057724.1

1 11 21 31 41 51
 | | | | |
 MAAARCWRPL LRGPRLSLET AANAATATE TTSQDVAATP VARYPPIVAS MTADSKAARL 60
 RRIERWQATV HAAESVDEKL RILTKMFMK YMVYPTFAL NADRWYQYFT KTVFLSGLFP 120
 PPAEPPEPE PEPEPALDLA ALRAVACDCL LQEHFYLRRL RRVERHYESE VISLPLDQL 180
 VSLVLGLLSP HNPALAAAL DYRCEVHFYW VRGEIIPRG HRRGRIDDLR YQIDDKPNMQ 240
 IRIKQLAEF VPLDYSVPIE IPTIRCKPDK LPLPKRQYEN HIFVGSKTAD PCYGHSTQFH 300
 LLPDKLRER LRRQNCADQI EVVFRANAIA SLPAWTGAQA MYQGFHSEAD VTRPFVSGAV 360
 ITDGKYFSFF CTQLNTLALT TQADQNNPRK NICWGTQSKP LYETIEDNDV KGFNDVLLQ 420
 IVHFLMRPK EBSKSLLEN

Seq ID NO: 81 DNA sequence

Nucleic Acid Accession #: FGENESE predicted

Coding sequence: 1..2070

1 11 21 31 41 51
 | | | | |
 ATGAGCGGTG CGGGGGTGGC GGTGGGAGC GGGCCCCCA GCTGCGCGAC CCCGGGCTCT 60
 CGGCGCGGCG GCCAGCGCCC CTCTGTGGGC GTCCAGTCTT TGAGGCGCGA GAGCCGCGAG 120
 CTGAGGCGAG GACAGCCGCA GAACCGGAAC CTGGAACCTG AGAAAGCCTT GCAGTCTCTG 180
 CAGCAGCAGC ACTCGGAGAT GCTGGCCAAAG CTCCATGAGG AGATCGAGCA TCTGAAGCGG 240
 GAAACCAAGG GTGAGCCGCG GCGGGGCCCT AGGCGCGGCC TGCTCCCCA GGCACACTCA 300
 ACACCTGCCG TCCCGCAGCA CAGAAACACA GCCATCAACT CCAGCACAGC CCGGGCTCA 360
 GGGGGAACAC AGGACCGGGA GCGCCCTCCAG ACTGTCTTTC CCCACCTGGC TGCACTGGGC 420
 OCTGTATGCC AACCCAGTGG GTACAGGTTT TGGGGGAOCT GACACAGATGC CGCTACCTCT 480
 AGCCGTGGCT GACAGATGTT ATGACGCCAA GCACAGCAGC TGCTGCTCTC GGGAGCCCA 540
 GGGCTCTAGG TCTATTGAGG GCGGCGGCTG GCCACAGGCT GCTCCCGAGA CTTCTCTCT 600
 CCAAGTAGAG CTGAATGGG AAGGAACCTC TGGGACAGCC CCGCCCTGCG TAGATCTTTG 660
 CCTCAGATTG CTGCTGTGGC CAGGCCAGAG ATTTCCAGCC CTATGGCTCT GAGTCTCTAC 720
 ATGCTGGGGG CCGAGGGGAT ATGACACAC TCCATCCAGG GATCCCTTCC TGCCATCTCG 780
 GCAGCAACCA TGGGGACAAA GGGAGGAAGC AGATCTCTGT TTCTTGCCA CTGTGCAAG 840
 GCMCTTCCC ATCTGACAG CGGCCCCCAG CCAGCCAGG ATCCTGGGCT GTGGTCTCAA 900
 GCTCACTTCC CATTATCTTT GGGCTGGGG CTGACATCAG GAGGACATCT GACTGGTGA 960
 TGGAGCCAGC CTGGGAACAT GCGAGCTGGG GCACTGCCTA GGGCTCTCCC TTCCAGGGA 1020
 GACATGAGGA AGGCGGTGTA GGGAGGCCCC TTCCCTAGCC GCTGTGGCAA CTTCACTGAG 1080
 CTGTCTGGG CAAAGTGTGG CCCAGTCTCG CAGCCCCAGC CCGTCACTGC TGGGAGCGCT 1140
 GACAGACAC GGGAGAGGCG CATGCTTTCC CTGCGGACCT GCTGTTCAT GTGTCCAAAG 1200
 CCCTCTGCT TTCCAGATGG CCCCTCAGGA AACCACCTTT CCAGGCGCTC TGCTCCCTTG 1260
 GCGCTCTGCT GGGTCTGCAT CAACGGAGTG TGGGTAGAGC GGGAGGAGCC CAGCCCTGCC 1320

5 AGGCTGAAGG AGGGCTCTCT ACGGACACAC AGGCCAGGAG GCAAGCGTGG GCGTCTTGGG 1380
 GCGCGTAGCG CCGACACTGT GCGCTCTCCT GCAGACAGCC TCTCCATGTC AAGCTTCCAG 1440
 TCTGTCAAGT CCATCTCTAA TTCAGCCAAC TCTCAAGGCA AGGCCAGGCC CCGACCCGGC 1500
 TCTTCAACA AGCAAGATTC AAAAGCTGAC GTCTCCAGCA AGCGCGGACT GGAAGAGGAG 1560
 CCCCCTACTT ACAACAGCAA GCTGGACAAA GTTCTCTGGG TACAAGGGCA GGCCAGAAAG 1620
 GAGAAAGCAG AGGCTCTTAA TGCAGGAGCT GCTGTATG GGAACAGCCA GCACCAGGGC 1680
 AGGCAGATGG GGGCGGGGGC ACACCCCCCA ATGATCTGTC CCTTCCCTCT GCGAAAGCCC 1740
 ACCACACTTA GGCAGTGGCA AGTGTCTATC CGGAGCTGT GGAATACCAA CCTCTGCAG 1800
 ACCAAGAGC TGCAGCAGCT CAAGTCCCTC CTGGAAGGGA GCCAGAGGCC CAGGCGAGCC 1860
 CCGGAGGAAG CTAGCTTTCC CAGGGACCAA GAAGCCACGC ATTTCCCCAA GGTCTCCACC 1920
 AAGAGCTCTT CCAAGAAATG CCTGAGCCCA CCTGTGGCG AGCGTGCAT CCTGCCGCA 1980
 CTGAAGCAGA CCCCAGAAAG CAACTTTGCC GAGAGGCAGA AGAGGCTGCA GGCAATGCAG 2040
 AAACGGCGCC TGCATCGCTC AGTGTCTTGA

15 Seq ID NO: 82 Protein sequence
 Protein Accession #: PGENESH predicted

20 1 11 21 31 41 51
 MSGAGVAAGT RPPSSFTPGS RRRRQRPSVG VQSLRPQSPQ LRQSDPQKRN LDLEKSLQFL 60
 QQQESEMIAK LHEETIEHLKR ENKGEFARGP RPALPQAHS TLPLPQHRNT AINSSTRIGS 120
 GGTQDGEPLQ TYLAHLAALA PVCQPSGYRF NGTWTDAATS SRGWIMLCISQ AQHVLLSGSP 180
 GPEVIAGROV ATGCSFDLPP PSRAEMGRNP WDSFPCPARSL PQIAAVARPR ISSPMALSPH 240
 25 MLGAQGLWTH SIQGSILPAIW AATMGTRGGS RVLPFCHLSK ALFHPDSEPH PAQDPGLWSQ 300
 AHFPLSLGLG LITSGHLTGG WSQPGNIAAG AVPRALPSQG DMEKGVGGGP FFSRCGNSSS 360
 LFWAKCGPSR QPQPCAGDA DRTREEMLS LGTCCSMCPK PSCFEDGPGG NHLGRASAPL 420
 GARHVCINGV WVEPGGSPSA RLKEGSSRTH RPPGKRGRJA GGSADTVRSF ADSLSMSSSQ 480
 SVKSIENSAN GCKARPPQPG SENKQDSKAD VSQKADLEES PLLHNSKLDK VPGVQGQARK 540
 EKAEASNAGA ACMGNSHQHG RQMGAAGHP MTLPLPLRKP TTLRQCSEVLI RELWNTLLQ 600
 30 TOELRHLKSL LKGSQRPQAA FEEASFPRDQ EATHFPKVST KSLSKKCLSP PVARRAILPA 660
 LRQTPKNMFA ERQKRLQAMQ KRRLHRSVL

35 Seq ID NO: 83 DNA sequence
 Nucleic Acid Accession #: NM_005264.1
 Coding sequence: 557..1954

40 1 11 21 31 41 51
 GAATTCGCGC CAGAAGAAAT CTGGCCTCGG AACACGCCAT TCTCCGCGCC GCTTCCAATA 60
 ACCACTAACA TCCTTAACGA GCATCCGAGC CGAGGGCTCT GCTCGGAAAT CGTCTGGCC 120
 CAACCTCGGC CTTCGAGCTC TCGAAGATTA CCGCATCTAT TTTTITTTTC TTTTITTTCT 180
 TTTOCTAGCG CAGATAAAGT GAGCCCGGAA AGCGAAGGAG GGGGCGGGGA CACCATTGCC 240
 CTGAAGAAGT AAATAAGTAA ATAAACAAAC TGGCTCTCG CCGCAGCTGG AGCGGTGCG 300
 45 TTGAGTCCAG GTTGGGTGGG ACCTGAACCC CTAAAGCGG AACCGCTCC CGCCCTCGCC 360
 ATCCCGAGCG TGAGTGGCGG CGGCGGTGG CTGCTGCCAG ACCCGAGTT TCTCTTTCA 420
 CTGGATGGAG CTGAACTTTG GCGCGCCAGA GCAGCACAGC TGTCCGGGA TOGCTGCAG 480
 CTGAGCTCCC TCGGCAAGAC CCAGCGGGCG CTCGGGATT TTTTGGGGG GCGGGGACCA 540
 GCCCGCGGCC GGCACCATGT TCCTGGCBAC CCGTCTCTTC GCGCTGCGC TCTTGGACTT 600
 50 GCTCCTGTGG GCGCAAGTGA GCGCGCGAGA CCGCTCTGAT TGGTGAAGS CCAGTGATCA 660
 GTGCTCTGAG GAGCAGAGCT GCAGCACCAA GTACCGCACG CTAGGCGAGT GCGTGGCGGG 720
 CAAGGAGACC AACTTCAGCC TGGCATCGCG CCGTGGAGCC AAGGATGAT GCGCGAGCGC 780
 CATGGAGGCC CTGAAGCAGA AGTCGCTCTA CAATGCGCG TGCAGCGGG GTATGAAGAA 840
 GGAGAAGAAC TTCTTGGCGA TTTACTGGAG CATGTACCAG AGCCTGCGAG GAAATGATCT 900
 55 GCTGGAGGAT TCCCCATATG AACCAGTTAA CAGCAGATTG TCAGATATAT TCGGGTGGT 960
 CCAATTCATA TCAGATGTTT TTCAAGCAAG GAGCAGCATT CCCAAGGGA ACAACTGCTT 1020
 GGATGCAGCG AAGGCTTGCA ACCTCGACGA CATTGCAAG AAGTACAGT CCGCTACAT 1080
 CACCCCGTGC ACCACAGCG TGTCCAACGA TGCTGCAAC CGCCGCAAGT GCCACAGGC 1140
 CCTCGCGCAG TCTCTTGACA ABGTCCGCG CAAGCACAGC TACGGAATGC TCTCTGCTC 1200
 60 CTGCGCGGAC ATGCGCTGCA CAGAGCGGAG GCGACAGACC ATCGTGCTG TGTGCTCCTA 1260
 TGAAGAGAGG GAGAGCCCA ACTGTTTGA TTTGCAAGG TCCTGCAAGA CCAATTACAT 1320
 CTGCGATCTT CCGCTGCGG ATTTTITTAC CAATGCGCG CAGAGTCAA GGTCTGTCAG 1380
 CAGCTGTCTA AAGGAAACT ACCTGACTG CCTCTGCTC TACTCGGGC TTATTGGCAC 1440
 AGTCATGACC CCAACTACA TAGACTCCAG TAGCCTCAGT GTGGCCCCAT GGTGTGACTG 1500
 65 CAGCAACAGT GGGAAAGACC TAGAAGAGTG CTTGAAATTT TTGAATTTCT TCAAGGACAA 1560
 TACATGTCTT AAAAATGCAA TTCAAGCCTT TGGCAATGGC TCCGATGTGA CCGTGGGCA 1620
 GGCAGCCTTC CAGTACAGA CCACCACTGC CACTACACC ACTGCCCTCC GGGTTAAGAA 1680
 CAAGCCCTCG GGGCCAGCAG GGTCTGAGAA TGAATTTCCC ACTCATGTTT TGCCACCGG 1740
 TGCARAATTA CAGGCACAGA AGCTGAAATC CAATGTGTGG GGCATACAC ACCTCTGTAT 1800
 70 TTCCAATGAT AATTATGAAA AAGAAGGTCT CGGTGCTTCC AGCCACATAA CCACAAAATC 1860
 AATGGCTGCT CCTCCAGCT GTGTCTGAG CCGCTGCTG GTCTGTGTGG TAACCGCTCT 1920
 GTCCACCTTA TTATCTTTAA CAGAACATC ATAGCTGCAT TAAAAAATA CAATATGGAC 1980
 ATGTAAAAAG TCAAAAACCA AGTTATCTGT TTCTGTCTT CTTGTATAGC TGAATTTCCA 2040
 GTTTAGGAGC TCAGTTGAGA AACAGTTCCA TTCACTGGA ACATTTTTTT TTTTCTTTT 2100
 75 AAGAAAGCTT CTGTGATGCC TTGCGGCTT CTGTGAAAAA CCTGATGCAG TGCTCCATCC 2160
 AAACCTCAAA GGCCTTGGGA TATGCTGTAT TTAAAGGGA CAGTTTGTAA CTTGGGCTGT 2220
 AAAGCAAACT GGGGCTGTGT TTTGATGAT GATGATCATC ATGATCATGA TGATTTTAA 2280
 AGTTTTACTT CTGCGCTTTC CTAGCTAGAG AAGGAGTTAA TATTTCTAAG GTAACCTCCA 2340
 TATCTCTTTT AATGACATTG ATTTCTAATG ATATAAATTT CAGCTACAT TGATGCCAAG 2400
 80 CTTTTTGGCC ACAGAAAGAA TTCTTACCAA GAGTGGGCTT TGTGGAACA GCTGTACTG 2460
 ATGTTCACTT TTATATATGT ACTAGCATTT TCCACGCTGA TGTITATGTA CTGTAAACAG 2520
 TTCTGCATCT TTGTACAAAA GAAAAACCA CCGGAATTC

Seq ID NO: 84 Protein sequence
 Protein Accession #: NM_005264.1

1 11 21 31 41 51
 5 MFLATLYFAL PLDLILLSAE VSGGDRLDV KASDQCLKEQ SCSTKYRTLK QCVAGKETNF 60
 SLASGLEAKD ECRSAMEALK QKSLYNCRCK RGMKKEKNCL RIYWSMYQSL QONDLLLEDSP 120
 YEPVNSRLSD IFRVVPFISD VFQOVHEHPK GNNCLDAKA CNLDDICKEY RSAYITPCPT 180
 SVSNDVCMRR KCHKALROFF DKVPAKHSYG MLFCSRDIA CTERRRQTIV PVCSEEREK 240
 PNCNLQDSC KTHYICRSRL ADFFTNCQPE SRSVSSCLKE NYADCLLAYS GLIGTVMTEN 300
 10 YIDSSSLVA PWDCNSNSGN DLEELCKFLN FFKDNTCLKN AIQAPGNSD VIVWQPPFPV 360
 QTTTATTTA LRWNKPLGP AGSENEIPTH VLPPCANLQA QKLKSNVSGN THLCISNGNY 420
 EKEGLGASSH IITKSMAPP SCGLSPLLVL VVTALSTLLS LTETS

Seq ID NO: 85 DNA sequence

Nucleic Acid Accession #: XM_027172.1

Coding sequence: 143..1405

1 11 21 31 41 51
 20 GGTGGAGACA CCGCCTCAGG GCTCGGTGCA CAGTGGACAT TTGGGGAGCG TTGTGGGTGA 60
 CCCCACACACA GGCACCTGGG ATGCAGGGGA GAGGGGGCCA AGGGGAAAG GGGCCAGAGT 120
 GTTGGCTTTG GATTCAGGAG GGATGGATTG CAGTCTAGC TTGCCACTTA TTAGGACTCC 180
 TGAGAGCAGC CTCCATGAGG CCTGGACCA GTGCATGACC GCTCTGACC TCTTCTCAC 240
 CAACAGTTTC TCAGAAGCAC TCAGCTACCT CAAGCCACGA ACCAAGGAAA GCATGTACCA 300
 CTCACCTGCA TATGCCACCA TCGTGGAGAT GCAGGCCATG ATGACCTTTG ACCCTCAGGA 360
 25 CATCTGTCTT GCGGCGAACA TGATGAAGGA GGCACAGATG CTGTGTGAGA GGCACCGGAG 420
 GAAGTCTTCT GTACAGATT CTTTCAGCAG CTTGTGTAA CCCCCACGC TGGGCCAATT 480
 CACTGAAGAG GAAATCCACG CTGAGGTCTG CTATGCAGAG TGCTGTCTGC AGGAGCAGC 540
 CCTGACCTTC CTGAGGGTT CCTCACACGG AGGGGCGATC AGGCCAGAG CTTGTGATGA 600
 TCCCTCTCAC GCTCTGAGCT GCCCACCCTG GCCAGGCCGT CAGCATCTTT TCTTCTGCA 660
 30 GAGCAGAGAG ATGTGTAGCT TCATCAAAGG CGGCATCAAA GTTGGAAACA GCTACAGAC 720
 CTACAAGGAG CTGAGACAGC TTGTTCAGTC CTCACAATAC TGCAGGGGTG AGAACCAACC 780
 GCACCTTTGA GAGGAGGTGA AGCTTGGTGT AGGGGCTTC AACCTGACAC TGTCCATGCT 840
 TCCTACTAGG ATCTGTAGGC TGTGGAGTT TGTGGGTTT TCAGGAACA AGGACTATGG 900
 GCTGTCTGAG CTGAGGAGGG GAGCCTCAGG GCACAGCTTC CGCTCTGTGC TCTGTGTCT 960
 35 GCTCTGCTG TGCTACCAACA CCTTCTCAC CTTCGTCTC GGTACTGGGA ACGTCAACAT 1020
 CGAGGAGGCC GAGAGCTCT TGAAGCCCTA CTGAACCGG TACCTAAGG GTGCCATCTT 1080
 CCTGTCTTT GACGGGAGGA TTGAAGTCAT TAAAGGCAAC ATTGATGAG TGAGTATGG 1140
 GGGTCCGGGC CGGGGCTGGG GATCCCTCAG GGTCTCCAG ACCAGCAGGA AGTCAGGCAC 1200
 40 ATGTGACATA CTCAGGGACA GGATAGACTG GGGGCGGGG GGGGGCAAG AGAGAACCAA 1260
 CCAGAGAGCA GGGCAGGAG AGGCCCTTCT GCAGAGCAG CCTGGGAAGA CAAGGGAGGA 1320
 GGAGGCATTG GTGTGCTCTG GGATTTTGAC TGGGAGATAT AGGACTGCAG CATGTCAGTG 1380
 GAGGAGGTG GAGGAGGTG CTGAGGGAG GCAGAGGTTA GGAAAGCCCA TCTGTTTAGG 1440
 GCATGACGAT TAGGCTGGAG TCIGGTACCT CCCCTCCATT ATAGCTCTCT CTTGCTCTTT 1500
 45 CATTTTGTTA CTAAAAACCA GAGTCTTAGG CGGGGGCTGT ATTTGAGCCC AACGTCTGT 1560
 AAGACTTAGG AGGTAAACCC AGGACTGGAG GCCAGATCTC CTGCTCTCTG GGGCCCAACC 1620
 TGAGCTTAGC ACAGGGCTGG ACCACTATGC CTTGAGGAG TCCCCTCTCT CTGTGTGTTT 1680
 GGGAGGTGG GAGGATCCAG AGGGGTTGGG GCTGGGTGGG CACCCCTCAG GCTGACCAGA 1740
 AGGTGCTTGC AGGCCATCCG GCGTTTCCAG GAGTGTCTGT AGGCCACGA GCACCTGGAAG 1800
 50 CAGTTCCACC ACATGTGCTA CTGGGAGCTG ATGTGGTCT TCACCTACAA GGGCCAGTGG 1860
 AAGATGTCTT ACTTCTAGCC CGACTGTCTC AGCAAGGAGA ACTGCTGTCT CAAGGTGGGG 1920
 TGATGCCAGC TGTTAGGGGC ATTGGGTGAC CAGGCTGAC TGTGTGCTC CAGACCACGG 1980
 GCCAATCCC TAACTGAACA CAGATGTCTC AGCTGGAATC TAAACATAAC CTTAAATCT 2040
 AATAGGACTC AGGCTTGGA GGAACATAAG ACCACAAGAG AACITCTGA ACCACAATGT 2100
 55 GTCAACAAGA GATTCTTAGT CACAAAGGAC AGAAACATGG CTCCTCTGT CCAGTAGAGT 2160
 TGTTCCTGAG TAGAAATGT TCTACTTATG CACTGCCCAA ATATGCTAGT CACCAAGCAC 2220
 ATGTGCACAA TGAGCACATG AAATGTGCTT AGTGCAACTG GGAACCTGAT TGTTTTCACT 2280
 TTTATTAAT TTTAATTAAC TAAATGTTAA ATTTAAATAG CCATGTAGGG CTTGTGGCCA 2340
 CTATATTGGA CTATGCAGGT CCAAAACACA AAAGGCTCAT ATAACTBAAC ATTCITGGCA 2400
 60 CATCCGACTT CAGGTAGGGC TGGATCCAGG AATTCAAATG ATGTGCTCTG GCTTGGTCTT 2460
 TCCATTTGTG GCTGTCTCT CTCCATGAC ATCTTTGTTT CTGCTGCATC TTGCGAGGAA 2520
 GGTCTCTTCC ATGTGACAGG CAAGGTGGCC ACCGGCTGCT TCTACTCATA TCCCTCCCTT 2580
 GGTTCACACC AGAGTCCAC AGTTTGAGCC TCACTGGTCT GACTTGTCTC CTGCCAATGC 2640
 CTGGAACAGA GGGGTGGGAA GATTCTCTAT GGGCTGAGAG TAGGAGAGGG GTGGTTCAC 2700
 65 AGAAGAAATG GATGCACCCA GAAATAGTGG AGGATTAAC AAGTGGCAT ACAGGCAGAA 2760
 CAAAGCCAAA CAGATGCCCG CTAACCAAG ATGAAATTTA TCATGTTAAG TATTGAAATA 2820
 AGTGTAGCT TGTACCATCA TAGTAATGAT AGTCAGAAA TTGGAACCAA GAGTCTTACA 2880
 ACCACCTAGC TCAGCAAAAG TCTAATCTGT TATTTGTAAT TACACAGGAC ATGTGTCTCT 2940
 70 ATGGCTTCAT CCCCCTGCCA TAGCAGACAT TGCTAATCAA TCCCTCTGCA TGAGCCTAGC 3000
 TGTGACCTTA GATACCTTCC CTGCAAGCT CCAGGCAGCC GTTAGAAGCT ACCCAGATTG 3060
 GCACCTGTTA GAGACCTGT CATCTGCATC TGCTCCCTCC CCTCTCTCTG CATGGAGGTC 3120
 TCTTCTGTAG CCGTGTCTGG GGAACCAAGT GAAGGGGAAT TCTCTGGGTC TGGGGCAGGT 3180
 GGGGCTGGGG AAGGGAGCAG GCTTCTACTG AGCTCTCAAC ATGTCTGGCA CCGTCTATG 3240
 75 CCTCCACACC ACACCCAGTG AGGAGAAGIT ATTATCTCCA TTGGGCACAT GAGGGCCACT 3300
 GATGCTAAGC GAGGACTGTG ACTTTCCAC CCACATGGCT GTTGGTGACA GACACCGGAT 3360
 TTGAGGACAG TCTGTGTGTA CTCAGAGGCC TCTGCTTCTC TCTGTCTCT AGAGTCTCTC 3420
 TGAGGAGTGG GGGCTTGGCC TGAGCCACCC CTGCTGTGTA AGGTGCTTCC TCAGGCCAGC 3480
 CTCCATAGGC CCGCATGCC CCTCTCATC ACCCTCTACT CCAAAAAGG ACAAGGCTCT 3540
 80 AGGGAACCTT TTTCTTTTT TTAGAGACAG GGTCTTGCTA TGTGGGTCAG GCTAGTCTTG 3600
 AACTCCTGGG CCCAAGCAAT CTTCCCACT CTGCTCTCTA AAGTGTCTGG ATTCTGGGCC 3660
 TTAGGGAACC TTTTGAACCT GAAAGTGACT CTCGAAGCTT TCTGTAGAG TGAGGTGGGT 3720
 GGCTGGGGCT AACCATATGG GGAAGGAGAG ACCTGGTGGG GGCACACAGC TGCTATATAG 3780
 AGGAACAGAA GGTGGGGCAG GCCTCCAGC TGTGAGGAAA TCTGGCTCAG TCCCGAGACC 3840
 GCTGTGTGGC TTTGGGGTGG CAGTCCCTC TCTCTCTGCA AACTCTCTCT CATGAGGACT 3900
 TCTGTGAAAA ATGGGGGTGG TAAACCTCTG AGTCTCTGGG CCCCACCAT TTTAATCTC 3960

5
10
15
20
25
30
35
40

CAGAACTTGG CCACCACAGA CTCCACCAAC TTCTAGTCCT GGGGCTGGG CCTCTGGCCA 4020
TTGCCATAGG CACCACCTGC TCTGTGCAGG CAGGCCCCC CTCTGCCAGG ATCCTCGAG 4080
GTCTAGCTGCT GGGTCTGACC CGCAGACCTT GGTGAGCCG CGGATGAACG GAGTATGCAG 4140
ACACAGGCTT GCTGTGCAGC AGATGGGGGA CCGCTGCCAG AGTCAGCAGC GGGCCCCATA 4200
AGCCTGCCAC GCTTGCATT ATTAGTACA GATGTAATGA CAAAGGCCCTA AAGCAAACTC 4260
CATTGTGGG TAATTAACAT TGTGCCGCC CCAGAAAGAG CAGTCTCTCG CATGATGATT 4320
AAAGGCCAGG TTCCGAGGCC TAAGTAAACC AACCTATCTA GATCAATTCC CTACTTCTT 4380
GTTATCTACT CTGAGAGAAT TCAGTGCCTT TCAGCCAAAT CCTTTCCGA AGCTTTTGCA 4440
AAACCTCCGA GCCTTCCAG GTTTGCTTCT TTCTGTAATT TTTCTCACC CCCTGACCTA 4500
TCTCTGCAG TCCAGCTGT GTAGGCCCTT GTGTTTCCC CAGTGTCTGG AGCCTAGAGG 4560
CTGAGATGGC CAGAAACAAG GTGGTGACAG TGGCGTCTC AGGGCTTGGG AAACCCAGG 4620
AGCTAAAGGC ATGCCAGGC AACCAAGAG GACAGGAAGG CTTCTGAGGA GAGACCTCTG 4680
AGGTGGGTCT TGAGAGAGAA GACTTAGGG AGGCAGAGT GAGGAAGTGA GAGGACACCC 4740
CAAGCCAAGA GGGCGGCAGG ACCAAGGCT CAGAGGCCAG GCGCTGCAG AGGGGCTGTG 4800
TGCCACAGG TGAGAGGTTT GTGTGGCAGA AGGCGAGGG GCTTGCATCA GGGGTGACAG 4860
CTGCTCTTTT GTCCAGCAT AGCCCTGTGA CATCCCTGGA GAGCTGGGGC GTCCCAACT 4920
CTAAGTCACA CCCCCATCC TAACCTTGT GTGTGAGTGA GGTGAGCTG TCTGTGGCA 4980
GAGGGGAAGA CTCTTGAGA TGAGCCTGT GAAGGGATAA TGGCATCCCG GGCAGAGGAG 5040
CAGCACAGGC AGAGGCTTGG GGAGAGTTTA AGGAGTGTAG GGGAGGAAAT GGCAGAAAGT 5100
GAGCCAGAAA TGGAGAGGGT AGGCGAGTC CTGGAGGACA TGAGTGGCTG TTTGGGCTT 5160
ATCCAGCAGT GGGGAGGCTT TGGCAGGCT GTGGCTTGA TAGGTGCTT AGAAAGCCCA 5220
CCAGCAGTTG CTGGGCCACC CCGCTGGCTG GGTCTGTTC TAAGGCAGGA AATCAAGCA 5280
TGAGCAGGAA AAGACCCCTT CAAGGCTCAC GTCTAGTGG GGAGACAGA AACACAGATG 5340
GGCAATATAA CACGATGTCT GGTTCAGTA AGTGCAGTGA AGAACAGCG AGGCTGGATG 5400
CAGGAGAGTA TGGAGAGGGC TTGTAAAGG GAGGTGGGG GAGGCTGTG TCAGAGGACA 5460
CCAGAAATGA GCGCAGAGC AGCAGGTGG AGTCAGTGG CAGGCGGTA GGCAGAGGG 5520
AGCTGGGCG GGCACAGAG GGCAGGAGT TGTGTGATGT TCTCTGGAA CCGCCCTGAG 5580
GCGTGTGTGT GGTGTGAGT CTGCAGGTGT CAGGAAAT GTAGGAGATG TCTCTGAGT 5640
GTGATGGAA ATAAACAGAT TTCCAGAGG AACTGACATG ATCTGACTTA AAAGSTCAG 5700
TGTGCAAAAT GCTTTCAGG GACAGGAGT GGGAGCAGG AGATAGGAGA CAATGTGTAC 5760
CAGGACAGCA GAAAGACAT CCGGTAGCC TGGACAGGG AGACGCTGT GAGATGGTGG 5820
CAGTCCGATA ATGAGAGCCG TAGGCCAAG CAGCAGGAT CCTAGAGTGA GACGGGAGGT 5880
AAAGTCACCG GACTTGGTG TCTCCAGTC AGGGCAGGG GAAAGGAGA GGACAGGGT 5940
GACCCGGAG GTTAAAGAT GAGCCGGGC CAGACGAGT GGCTCATGCC TGTATCTTA 6000
GCATTTGGG ATGGLQLEB GGCAGATAG CTGAGGTCA GAGATTGAA ACCAGCTGG 6060
CCACATGTGT GAAACCCGT CTCTACTAA ATATACAAA ATTAGCCTGG CBTGTGTGTG 6120
CATGCTTGA GTCCAGCTA TTCAGGAGG TGAGGCAACA AGAATCGCTT GAACCTGGG 6180
GGGGAGGTT GCAGTAGGC GAGATCGCG CATAGCACTC CAGCCTTAGC CTGGGCGACA 6240
GAGCGAGACC ACATC

Seq ID NO: 86 Protein sequence
Protein Accession #: XP_027172.1

45
50
55

1 11 21 31 41 51
MDSPSLPLI RTPESLLEHA LDQCHTALDL FLTNQFSEAL SYLKPRTKES MYESLTYATI 60
LEMQAMTFD PQDILLAGNM MKEAQMQLQR HRRKSSVIDS PSLVNRPTL GQTEERIEA 120
EVCYAECLLQ RAALTFLQGS SHGGAVRERA LHDPSEACSC PFGPGRQLHF LLQDENMVSF 180
IKGKIKVENS YQTYKELDSL VQSQYQKGE NHPEFEGGVK LGVGFNLTSL SMLPTRILRL 240
LEPVGFSGNK DYGILLQLEB ASGHSFRSLV CVMILLCYHT PLTFVLGTGN VNIEAEKLL 300
KPLNRYEPKG ALFLFPAGRI EVIKENTDAV SDGSPGRGWG SLGVSTSRK SGTCDILLDR 360
IDWGRGGGQE RTNQRAGAGH ALLAEQPKGT REBEAFVVPG ILTGRYRTAA LQWREVEGGA

Seq ID NO: 87 DNA sequence
Nucleic Acid Accession #: AB007921
Coding sequence: 143..1363

60
65
70
75
80

1 11 21 31 41 51
GGTGGAGACA CCGCTCAGG GCTCGGTGCA CAGTGGACAT TTGGGGAGCB TTGTGGGTGA 60
CCCCACACA GGCAGTGGGA ATGCAGGGGA GAGGGGGCCA AGGGGGAAG GGGCCAGAT 120
GTGTGGCTTG GATTTCAGGAG GGTATGATTC CAGTCTTAGC TTGCCACTTA TTAGGACTCC 180
TGAGAGCAGC CTCCATGAGG CCGTGGACCA GTGCATGACC GCGCTGGACC TCTTCTTAC 240
CAACCACTTC TCAGAGGAC TCAGCTACCT CAAGCCGAGA ACCAGGAGAA GCATGTACCA 300
CTCACTGACA TATGCCACCA TCTTGGAGAT GCAGGCCATG ATGACCTTTG ACCCTCAGGA 360
CATCTGTCTT GCGGGCAACA TGATGAAGGA GGCACAGATG CTGTGTGAGA GGCACCGAG 420
GAAGTCTTCT GTAAAGATT CTTTACGAC CTTGGTGAAC GCGCCACCG TGGGCCAATT 480
CACTGAAGAG GAAATCCACG CTGAGGTCTG CTATGCAGAG TGCTTGTCTG AGGAGCCAGC 540
CTGACCTTC CTGAGGGTT CTTACACGG AGGGGCAGTC AGGCCAGAG CCTTGCATGA 600
TCCCTCTCAC GCTTGCAGCT GCGCACCTGG GCGAGGCGCT CAGCATCTTT TCTCTCTGCA 660
GGAAGAGAAC ATGCTGAGCT TCATCAAGG CCGCATDAAA GTTCGAAACA GCTACCABAC 720
CTCAAGAGAG CTGACAGGCC TTGTTTCACT CTCACAATAC TGCAAGGGTG AGAACCAACC 780
GCATTTTGAA GGAGGAGTGA AGCTTGGTGT AGGGGCTTC AACCTGACAC TGTCCATGCT 840
TCTTACTAGG ATCTGTAGGC TGTGTGAGTT TGTGGGGTTT TCAGGAAACA AGGACTATGG 900
GCTGTCTGAG CTGAGGAGG GAGCGTCAGG GCACAGCTTC CGCTCTGTGC TCTGTGTAT 960
GCTCTGTGCT TGCTACACA CTTTCTTAC CTTCGTCTC GGTACTGGGA AGTCAACAT 1020
CGAGGAGGCC GAGAGCTCT TGAAGCCCTA CTTGAACCGG TACCCTAAGG GTGCCATCTT 1080
CCTGTTCTTT GCAGGAGGA TTGAAGTCAT TAAAGGCAAC ATTGATGCGG TGAGTGTAGG 1140
GGGTCCGGGC GGGGCTGGG GATCCCTGG GGTCTCCAG ACCAGCAGGA AGTCAGGCAC 1200
ATGTGACATA TCCAGGACA GGATAGACTG GGGGCGGGG GGGGGGCCAA GAGAGAACA 1260
ACCAAGAGAG AGGGGCGAGA GAGGCCCTTC TGGCAGAGCA GCCTGGGAG ACAAGGGAGG 1320
AGGAGCAATT TGTGCTGCTT GGGATTTTGA CTGGGAGATA TAGGACTGCA GCATTGCACT 1380
GGAGGAGAGT GGAGGAGGT GCTTGAAGGA GGCAGAGGTT AGGAAAGCCC ATCTGTCTAG 1440
GTCATGACGA TTAGGCTGGA GTCTGTATAC TCCCTTCCAT TATAGCTCTC TCTGTCTCTT 1500

	TCATTTTGT	ACTAAAAACC	AGAGTCTTAG	GCGGGGGCTG	TATTTGAGCC	CAAGTCATG	1560
	TAGAGCTTAG	GAGGTAAACC	CAGGACTTGA	GGCCAGATCT	CCTGGCTCCT	GGGGCCCCAC	1620
	CTGAGCCTAG	CACAGGGCTG	GACCACTATG	CCCTGGAGGA	GTCCCGGTCT	GCTGTGGTGT	1680
5	TGGGAGGTTC	GGAGGATGCA	GAGGGGTGG	GGCTGGGTGG	GCACCCGTCA	GGCTGACCAG	1740
	AAGGTGCTTG	CAGGCCATCC	GGCGTTTCGA	GGAGTGTCTG	GAGGCCCCAG	AGCCTGGAA	1800
	GCAGTTCCAC	CACATGTGCT	ACTGGGAGCT	GATGTGGTGC	TTCACCTACA	AGGGCCAGTG	1860
	GAAGATGTCC	TACTTCTACG	CCGACCTGCT	CAGCAAGGAG	AATGTCTGGT	CCAAGGTGGG	1920
	CTGATGCCAC	GTGTTAGGGG	CATTGGGTGA	CCAGGGCTGA	CTGTGTGCC	CCAGACCAGG	1980
10	GGCCAAATCC	CTAATGAAAC	ACAGATGTCT	CAGCTGGAAAT	CTAAACATAA	CCTTAAATTC	2040
	TAATAGGACT	CAGGCTTGGG	AGGAACTAAA	GAOCCACAGA	GAAACTTCTG	AACCAATATG	2100
	TGTCAAAAG	AGATTTCTAG	TCACAAAGGA	CAGAAACATG	GCTCCCTCTG	TCCAGTAGAA	2160
	CTGTTCCTGT	GGTAGAAATG	TCTACTTAT	GCAGTCCCCA	AATATGGTAG	TCACCAGCCA	2220
	CATGTGCACT	CTGAGCACAT	GAAATGTGCC	TAGTGCACAT	GGGAACTGAA	TGTGTTTCAG	2280
15	TTTTATTATA	TTTTAATTA	CTAAATGTA	AATTTAAATA	GCCATGTAGG	GCTTGTGGCC	2340
	ACTATATTGG	ACTGCTTGGG	TCCAAACAC	AAAAGGCTCA	TATAACTGAA	CATTCTTGCC	2400
	ACATCCGACT	TCAGGTAGGG	CTGGATCCAG	GAATTCAAAT	GATGTGCTCT	GGCTTGGTCT	2460
	TTCATTTGT	GGCTGTGCTC	TCTCTATGA	CATCTTTGTT	TCTGTGCAAT	CTTTCAGGA	2520
	AGGTCTCTC	CATGTGACAG	GCAAGGTGGC	CACGGGCTGC	TCTACTCAT	ATCTCCCTCT	2580
20	TGTTTCAAC	CAGAGTCCCA	CAGTTTGAGC	CTCACTGGTC	TGACTTGTCA	CCTGCCAATG	2640
	CCTGGAACAG	AGGGCTGGGA	AGATTCTTCA	TGGGCTGAGA	GTAGGAGAGG	GGTGGTTCCA	2700
	CAGAAGAAA	TGATGCAACC	AGAAATAGTG	GAGGGATTAA	CAAGATGCCA	TACAGCAAA	2760
	ACAAAGCCAA	ACAGATGCCC	GCCTACCAAA	GATGAATTT	ATCATGGTAA	GTATTGAAT	2820
	AAGTGTAGC	TGTGACCATC	ATAGTAATGA	TAGTGCAGAA	ATTGGAACCA	AGAGTCTTAC	2880
25	AACCACTTAG	CTCAGCAAAC	GTCTAATCTG	TTATTGTATA	ATACACAGGA	CATGTGTCTT	2940
	CATGCTTCA	TCCCTTGCCC	AUGGCAGACA	TGCTAATCA	ATCCTCTGCC	ATGAGCCTAG	3000
	CTGTGACCTT	AGATACCTTC	CTGCAAAAGC	TCCAGGCAGC	CGTTAGAACT	GACCTACATT	3060
	GGCACCCTGT	AGGAGACCTG	TCATCTGCAT	CTGGTCCCTC	CCCTCTCTCT	GCATGAGGT	3120
	CTCTTCCGTA	TCTGTGCTGC	GGGACCAAGC	TGAAGGGGAA	TTCTCTGGT	CTGGGGCAGG	3180
30	TGGGGCTGGG	GAGGGGAGCA	GGCTTCTACT	GAGCTCTCAA	CATGTCTGGC	ACCTGTCTATG	3240
	TCTTCCACAC	CACACCCAGT	GAGGAGAAGT	TATTATCTCC	ATTGGGCACA	TGAGGGCCAC	3300
	TGATGCTAAG	CGAGGACTGT	GACTTTCCCA	CCACATGGC	TGTTGGTGAC	AGACACCGGA	3360
	TTTGAGGACA	GTCTGTGTGT	ACTCAGAGGC	CTCTGCTCT	CTCTTGCTCC	TAGAGTCTCT	3420
	CTGAGGAGTC	GGGCTTGCC	CTGAGCCAC	CCGCTGTGTC	AAGGTGCTTC	CTCAGGCCCA	3480
35	GCTCCCATGG	CCCCACACCC	CCCTCCTCAT	CACCTCCTAC	TCCCAAAAG	GACAAAGCCT	3540
	CAGGGAACCT	TTTTCTTCTT	TTTAGAGACG	GGGTCTGCT	ATGTTGGTCA	GGCTGGTCTT	3600
	GAACCTCTGG	GGCCAGGCAA	TCTTCCGCCC	TCTGCTCCT	AAAGTGTGCG	GAATCTGGGC	3660
	CTTAGGGAAC	CTTTTGTGAC	TGAAGGTGAC	GCTCGAAGCC	TTCCTGTAGA	GTGAGGTGGG	3720
	TGGCTGGGCG	TAAACCATATG	GGGAAGGAGA	GACCTGGTGG	GGGCACACAG	CTGCTATATA	3780
40	GAGGAACAGA	AGGTGGGCCA	GGCTCCCGAG	CTGTCAAGAA	ATCTGGCTCA	GTCCCCAGAC	3840
	CGCTGTGTGG	CTTGTGGGTG	GCCAGTCCCT	CTCTCTGCG	AACACTCTCC	TGATGAGGAC	3900
	TTCTGTGAAA	AATGGGGGTT	GTAACTCTCT	GAGTTCTGGG	GGCCCAACCA	TTTCTAATCT	3960
	CCAGAACTTG	GGCCACACAG	ACTCCACCAA	CTTCTAGTCC	TGGGGCTTGG	GGCTCTGGCC	4020
	ATTGCCATAG	GCACCATAG	CTCTGTGTCAG	GCAGGCCCTC	CCCTCTGCCAG	GATCCTCCGA	4080
45	GGTCAGCTGC	TGGGTCTGAC	CCGCAGACCC	TGGCTGAGCG	ACGGATGAAC	GGAGTATGCA	4140
	GACACAGGCT	TCTTGTGTCAG	CAGATGGGGG	ACCCCTGCCA	GAGTCAGCAG	CGGCCOCCAT	4200
	AAGCCTGCCA	CGCTTGCATT	TATTAGTAC	AGATGTAAATG	ACAAAGGCCCT	AAAGCAAACT	4260
	CCATTGTGTG	GTAAATTAACA	TTGTGCCCCC	CCAGAAAGA	GCAGTCTCTC	GCATGATGAT	4320
	TAAAGGCCAG	GTTCOGAGGC	CTAAGTAAAC	CAACTTATCT	AGATCAATTC	CCCTACTTCT	4380
50	TGTTATCTAC	TCTGAGAGAA	TTCACTGCTC	TTCAAGCCAA	TCTTTTCCCG	AAGCTTTTGC	4440
	AAAACTCTCG	AGCCTTCCAA	GGTTTGTCTC	TTTCTGTAAT	TTTTCTCACC	ACCCCTGACT	4500
	ATCTCTGTCA	GTCAAGCCCTG	TGGAGGCTTT	TGTGTTTCCC	CCAGTGTCTG	CAGCTAGAG	4560
	GCTGAGATGG	CCAGAAACAA	GGTGGTGACA	GTGGCTGTCT	CAGGGCTTGG	GAAACCCAG	4620
55	GAGCTAAAGG	CATGCCCCAG	CAACCAAGA	GGACAGGAAG	GCTTCTGAGG	AGAGACCTCT	4680
	GAGGTGGGCT	TTGAGAGAGA	AGGACTTAGG	GAGGCAGAGT	GGAGGAAGTG	AGAGGACACC	4740
	CCAGGCCAAG	AGGGCGGCGAG	GACCAAGGCT	TCAGAAAGCC	GGGCGCTGCA	GAGGGGCTGT	4800
	GTGCCACAGG	GTGAAGAGTT	TGTGTGGCAG	AAGGGCAGGG	GGCTTGCTATC	AGGGGTGACA	4860
	GCTGCTCTTT	TGTCCACAGA	TAGCCCCCTGT	ACATCCCTGG	AGAGCTGGGG	GTCCCAACAC	4920
	TCTAAGTCAAT	AGCCCCATC	CTAACCTCTG	TGGTGCAGTG	AGGGTGAGCT	GTCTGTGGGC	4980
60	AGGAGGCAAG	ACTCTTGGAG	ATGAGCCCTG	TGAAGGATA	ATGGCATCCC	GGGCCAGGGA	5040
	CGAGCACAGG	CAGAGGCCCTG	GGGAGAGTTT	ANGGAGTGT	GGGAGGAAA	TGGCAGAGGA	5100
	TGAGCCAGAA	AAAGAAAGGT	TAGGGCAGGT	CCTGGAGGAC	ATGAGTGGCT	GTTTGGGCTT	5160
	TATCCAGCAG	TGGGGAGGCC	TGGCAGGCT	TGTGGCTTAG	ATAGGTGCTT	TAGAAAGCCC	5220
65	ACCAGCTGAT	GCTGGGCCAC	CCCGCTGGCT	GGGTCTGTT	CTAAGGCAGG	AAATACAGGA	5280
	ATGAGCAGGA	AAAGACCCCC	TCAAGGCTCA	CGTCTTAGTG	GGGAGACAG	AAACACAGAT	5340
	GGGCAATATA	ACAGCATGTC	TGGTTCCAGT	AAGTGCACTG	AAGAACAAGC	GAGGCTGGAT	5400
	GCAGGAGTGT	ATGGGAGGGG	CTTTGTAAAG	GGAGGTGGGG	GGAGCCTGT	CTCAGAGGAC	5460
	ACCAGAAATG	AGCGCAGGAG	CAGCAAGCTG	CAGTCAATG	GCAGGCCCTT	AGGGCAGAGG	5520
70	GAGCTGGGCA	GGGCACAGCA	GGGCAGGAGT	GTGTTTGATG	TGTCTTGGGA	ACCGCCCTGA	5580
	GGCCGCTGTG	TGGCTGGAGT	GCTGCAAGTG	TCAAGGAAAT	TGTAGGAGAT	GTCTCTTGAG	5640
	TGTGATGAAA	TATAACCAAG	TTTCCAGAG	GAAGTACAT	GATCTGACTT	AAAAAGGTCA	5700
	GTGTGCGAAA	TGGCTTGCAG	GGGACAGGAG	TGGAGCAGG	GAGATAGGAG	ACAATGTGTA	5760
	CCAGGACAGC	AGAAAGACAT	CCCGGTAGC	CTGGAACAGG	GAGACGCTGT	GGAGATGGTG	5820
	GCAGTCCOAT	AATGAGAGCC	GTAGGGCAAG	GCCAGCAGGA	TCTAGAGTG	AGACGGGAGG	5880
75	TAAAGTCAAC	GGGACTTGGT	GTCTCCACGT	CAGGGCCAGG	GGAAAGGGAG	AGGACAGGGG	5940
	TGACCCGGGA	GGTTAAAGAT	GGGACCGGGG	CCAGACGAG	TGGCTCATCG	CTGTAACTCT	6000
	AGCACTTTGG	GAGGCTGAGG	CGGGCGGRTG	GCTTGAGGTC	AGGAGTTTGA	AACCGGCTTG	6060
	GCCACATATG	TGAACCCCCG	TCTCTACTAA	AATATACAAA	AATTAGCTTG	GGTGGTGGT	6120
	GCATGCCCTG	AGTCCAGCT	ATTCAAGAGG	CTAGGGCCAG	AAGATCGCT	TGAACCTGGG	6180
80	AGGCGGAGGT	TGCATGAGC	CGAGATCGCG	CCATAGCTAT	CCAGCCTTAG	CCTGGGCGAC	6240
	AGAGCGAGAC	CACATC					

Seq ID NO: 88 Protein sequence
Protein Accession #: BAA32297.1

```

1      11      21      31      41      51
5  VETPPQGSVH SCHLGSVVGD PHITGTGNAGE RSPRGKQARV LALDSGGMDS SPSPFLIRTP 60
    BSSLREALDQ CMTALDLFLT NQFSEALSXL KPRTKESMYH SLTYATILEM QAMMTFDPQD 120
    ILLAGNMMEK AQMLCQRHRR KSSVTDSPSS LVNRPITLQGF TEEIHAIEVC YAECLLQRAA 180
    LTFLOGSSHG GAVRPRALHD PSHACSCPPG PGRQHLFLLO DENMVSPFKG GIKVRNSYQT 240
    YKELDSLIVQS SQYCKGENHP HPEGGVKLGV GAFNLTLNML PTRILRLLEF VGFSGNKDYG 300
    LLQLEEGASG HSFRSVLCVM LLLCYHTFLT FVLGTGNVNI EEAEKLLKPY LNRYPKGAIF 360
10 LFFAGRIEVI KGNIDAVSDG GPRGRWNSLG VBQTSRKSQT CDILDRIDW GRGGGFRENQ 420
    PESRRRRGFS GRAAWEDKGG GGICGAWDFD WEI

```

Seq ID NO: 89 DNA sequence

Nucleic Acid Accession #: AF007170

Coding sequence: 73..1725

```

1      11      21      31      41      51
20 AAGGAGGCGG CCTCCGGGAA AAGCGACCGC AGGACTCCTG AGAGCAGCCT CCATGAGGCC 60
    CTGGACCACT GCATGACCGC CCTGGACCTC TTCTCAACCA ACCAGTTCTC AGAAGCACTC 120
    AGCTACCTCA AGCCGAGAAC CAAGGAAAGC ATGTACCACCT CACTGACATA TGCCACCATC 180
    CTGGAGATGC AGGCCATGAT GACCTTTGAC CCTCAGGACA TCCTGCTTGC OGGCAACATG 240
    ATGAGAGGAG CACAGATGCT GTGTCAAGAG CACCGGAGGA AGTCTTCTGT AACAGATTCC 300
    TTCAGCAGCC TGCTGAACCG CCCCAAGCTG GGCCTATTCA CTGAAGAAGA AATCCACGCT 360
25 GAGGTCTGCT ATGCAGAGTG CTTGCTGAG CGAGCAGGCC TGACCTTCTT GCAGGACGAG 420
    AACATGGTGA GCTTCATCAA AGGCGGCATC AAGTTCGAA ACAGCTACCA GACCTACAG 480
    GAGCTGGACA GCTTGTGTCA GTCTTCACAA TACTGCAAG GTGAGAACCA CCGCACTTT 540
    GAAGGAGGAG TGAAGCTTGG TGTAGGGGCC TTCAACCTGA CACTGTCCAT GCTTCTTACT 600
    AGGATCCTGA GCTGTGTGGA GTTTGTGGGG TTTTCAGGAA ACAAGGACTA TGGGCTGCTG 660
    CAGCTGGAGG AGGGAGCGTC AGGCCACAGC TTCCGCTCTG TGCTCTGTGT CATGCTCCTG 720
    CTGTGCTACC ACACCTTCTT CACCTTCGTG CTGGTACTTG GGAAGCTCAA CATCGAGGAG 780
    GCGGAGAGC TCTTGAAGCC CTACCTGAAC CGGTACCTTA AGGCTGCCAT CTTCTGTTC 840
    TTTGAGGGGA GATTGAAGT CATTAAAGGC AACATTGATG CAGCCATCCG GCGTTTCGAG 900
    GAGTGTCTGT AGGCCACGCA GCACTGGAAG CAGTTTCAAC ACATGTGCTA CTGGGAGCTG 960
35 ATGTGTGTCT TCACTACCAA GGGCCAGTGG AAGATGTCTT ACTTCTACGC CGACCTGCTC 1020
    AGCAGGAGGA ACTGTGGTTC CAAGCCACAC TACATTTACA TGAAGGCCGC CTACCTCAGC 1080
    ATGTTTGGCA AGGAGGACCA CAAGCCGTTT GGGGACGAG AAGTGGAAIT ATTTGAGCT 1140
    GTGCCAGGCC TGAAGCTCAA GATTGCTGGG AAATCTCTAC CCACAGAGAA GTTTGCCATC 1200
    CGGAAGTCCC GCGCTACTTT CTCTCTCAAC CCTATCTGCG TGCCAGTGCC TGCTCTGAG 1260
    ATGATGTACA TCTGGAACCG CTACGCCGTG ATTGGGAAGC AGCCGAAACT CACGGATGGG 1320
    ATACTTGAGA TTATCTATAA GGTGGAAGAG ATGCTGGAGA AAGGCCGAGA GAACGAGTAC 1380
    TCAGTGGATG ACGAGTGCTT GGTGAATTTG TTGAAGGCC TGTGTCTGAA ATACCTGGG 1440
    CGTGTCCAGG AGGCCGAGGA GAATTTTAGG AGCATCTCTG CCAATGAAAA GAAGATTAAA 1500
    TATGACCACT ACTTGATCCC AAACGCCCTG CTGGAGCTGG CCTGCTGCTT TATGGAGCAA 1560
45 GACAGAAACG TCTGGAACCG CAACCTTTTG GAATCTGCCA AGCAAAACTA CAAGAATTAC 1620
    TCCATGGAGT CAAGGACACA CTTTGAATC CAGGCAGCCA CACTCCAAGC CAAGTCTTCC 1680
    CTAGAGAACA GCAGCAGATC CATGCTTCCA TCAGTGTCTT TGTAGCTTTG TGCAGCAGTT 1740
    CCGGGCTGGA AGACAGAGAC AGCTGGACAG AGCTCTTGAA AACATTTCAA AATACCCCTT 1800
    CCCCCTGCCC TGCCCTGCTT TTGGGGTCCA CCGGCACTCC AGTTGGATGG CACAACATAG 1860
50 TGTATCCGTT CAGGAAGCCG GCTGGCATTT TCACCACTGT AGCCAGAGGC CTTTGGCAGG 1920
    GGCAGAGCAG GTGAGAGCCT CTGCTGCCC TATCACACAT ACGGTACTIT GCTTTTCACT 1980
    GTGATGTTTA AGAGAAATGA TGAACATTT ACATTTTCTT TAGAATAACA TTGATGGGAT 2040
    CACAGTTGGC TTAAAAAACC AACAAACATC AACCACTGT AAGTCTTTGT CTTCACTCAT 2100
    TATCATCTGG AGTTAAATCT CTTTATATGA TGAAGCCAAA GGCCTAAATG CTTTCTAAAT 2160
55 TCAGCAAGTT CTAAGCTTGT GTGACGGAAG GTCTCTCAGA GGACCTGAGG AATGCTTGGG 2220
    AGAGGCTAAG CCTCAGGCTT CAATGCTTCT GGGGTTGGGC ATGAGGATGT ACACAGACAC 2280
    CCCTACCTTT ACTACTCACA CTTCATTTC CTCTTTTGT AAATTTCCAA TTTAAAAATC 2340
    AAGCAAGTCT TTTTAGTGG ATAAAAATCT AGCTCTTCTG TAGAAAAATC AATCTCTACC 2400
    AGTAGAAAT GGCAGGCTT GATGGAAGAG CTGTGTAGCC CTTTCTATGC CAAGGCCAGG 2460
60 AAATTTGGGG GGCAGGAGGA GGTCTTCAGA ATCCAGTCTG TATCTTTGCT GTATGCCAAA 2520
    CTGAAACCAC TGGGAATAAT TTATGAAACA TAAAAATCT CTGTACTTCA CTCCAAGGTA 2580
    CATTGTCTTA CTGACAGCAT TTTTGTAAAA ACTGTTATTC TTGAAAAAAA AAAAAAAA 2640
    AA

```

Seq ID NO: 90 Protein sequence

Protein Accession #: AAC39582.1

```

1      11      21      31      41      51
70 KRAASGKSDR RTPSSSLHBA LDQCMTALDL FLTNQFSEAL SYLKPRTKES MYHSLTYATI 60
    LEMQAMMTED PQDILLAGNM MKBAQMLCQR HRRKSSVIDS FSSLVNRPIL GQFTBEEIHA 120
    EVCYAECLLQ BAALITFLQDE NMVSPFKGKI KVRNSYQTYK ELDSLIVQSSQ YCKGENHPHF 180
    EGGVKLGVA G FNLTLNMLPT RILRLLEFVG FSGNKDYGLL QLEEGASGHS FRSVLCVMLL 240
    LCYHTFLTFF LGTCNVNTEE AEKLLKPYLN RYPKGAIFLF FAGRIEVIK NIDAAIRRPE 300
75 ECCEAQQHWK QFHEMCYHEL MWCFTYKQW KMSYFYADLL SKENCWSKAT YIYMKAAYLS 360
    MFGKEDHXPV GDDVELEFRA VPKLKLKLAG KSLPTKFAI RKSRRYFSSN PISLFPVPALE 420
    MYIINNGYAV IGKQPKLTDG ILELITKAE MLERGFENEY SVDECLVKL LKGLCLKYLK 480
    RVQEAENRFR SISANEKKIK YDHYLIPNAL LELALLLMEQ DRNEBAIKLL ESAXQNYKNY 540
    SMESRTHFRI QAATLQAKSS LENSRRSMVS SVSL

```

Seq ID NO: 91 DNA sequence

Nucleic Acid Accession #: Bos sequence

Coding sequence: 31..906

1	11	21	31	41	51	
CGGGTCGACC	CACGCGTCCG	GGGAGAAAGG	ATGGCCGGCC	TGGCGGCGCG	GTTGGTCCTG	60
CTAGCTGGGG	CAGCGGCGCT	GGGAGGCGGC	TCCAGGGGCG	ACCGTGAGCC	GGTGTACCGC	120
GACTCGGTAC	TGCAGTGGGA	AGAGCAGAAC	TGCTCTGGGG	GCGCTCTGAA	TCACTTCCGC	180
TCCCGCCAGC	CAATCTACAT	GAGTCTAGCA	GGCTGGACCT	GTGGGGACGA	CTGTAAATAT	240
GAGTGTATGT	GGGTACACBT	TGGGCTCTAC	CTCCAGGAAG	GTCACAAAGT	GGCTCAGTTC	300
CATGGCAAGT	GGCCCTTCTC	CGGTTCTCTG	TTCTTTCAAG	AGCCGGGCATC	GGCGTGGGCC	360
TGCTTTCTCA	ATGGGCTGGC	CAGCTTGGTG	ATGCTCTGCC	GCTACCGCAC	CTTGTGTCCA	420
GGCTCCTCCC	GCATGTATCA	CACCTGTGTG	GGCTTGGGCT	GGGTGTCCCT	CAATGCATGG	480
TTCTGGTCCA	CAGTGTTCAC	CACCAGGGAC	ACTGACCTCA	CAGAGAAAAT	GGACTACTTC	540
TGTGCTTCCA	CTGTATCCTT	ACACTCAATC	TACCTGTGCT	GGGTCAAGCT	CATCCGCTTC	600
GACTATGGCT	ACAACCTGGT	GGCCAACGTG	GCTATTGGCC	TGGTCAAGCT	GGTGTGGTGG	660
CTGGCCCTGT	GGCTGTGGAA	CCAGCGGGCG	CTGGCTCACG	TGGCGAAGTG	CGTGGTGGTG	720
GTCTTGTCTG	TCCAGGGGCT	GTCCCTGCTC	GAGCTGCTTG	ACTTCCACCC	GCTCTTCTGG	780
GTCTTGTCTG	CCCATGGCCT	CTGGCACATC	AGCACCATCC	CTGTCCACGT	CCCTTTTCTC	840
AGCTTTCTGG	AAGATGACAG	CTGTACCTCG	CTGAAGGAAT	CAGAGGACAA	GTTCAAGCTG	900
GACTGAAGAC	CTTGGAGCGA	GTCTGCCCCA	GTGGGGATCC	TGCCCCCGCC	CTGCTGGCCT	960
CCCTTTCTCC	CTCAACCTTT	GAGATGATTT	TCTCTTTTCA	ACTTCTTGAA	CTTGGACATG	1020
AAGGATGTGG	GCACGAAATC	ATGTGGCCAG	CCCAACCCCT	GTGGCCCTTC	ACCAGCCTTG	1080
GAGTCTGTTC	TAGGGAAGGC	CTCCAGGCAT	CTGGGACTCG	AGAGTGGGCA	GGCCCTCTAC	1140
CTCTGTGAGC	TGAACTGGGG	TGGAACTGAG	TGTGCTCTTA	GCTCTACCGG	GAGGACAGCT	1200
GGCTGTCTTC	TCCCATCTAG	CTCTCTCCCC	ACATCCCCAG	CTGCTGGGCT	GGGTCTCTAA	1260
GGCTCTGTCT	TACCTGGGAG	ACCAGGGACC	ACAGGCCCTTA	GGGATACAGG	GGGTCCCCCT	1320
CTGTATCCAC	CCCCACCCCT	CTCCAGGAC	ACCACTAGGT	GGTGTGGTAT	GCTGTCTCTT	1380
TGGCCAGCCA	AGGTTTCAAG	CGATTTCTCC	CATGGGATCT	TGAGGGACCA	AGCTGTCTGG	1440
ATTGGGAAGT	AGTTTCAACC	TGACCTTTGC	CTTACCCAGG	TTCCACAGGAG	GGCTCACCAT	1500
ACTCCCTTTC	AGGCCCAGGG	CTCCAGCAAG	CCCAGGGCAA	GGATCTGTGT	CTGCTGTCTG	1560
GTGTGAGGCC	TGCCACCGTG	TGTCGGGAGT	GTGGGCCAGG	CTGAGTGCAT	AGGTGACAGG	1620
GGCTGTAGCA	TGGGCTGTGG	TGTGTGTGAG	CTCAGGCACT	AGGTGCGCAG	TGTGGAGACG	1680
GGTGTGTCTG	GGGAGAGAGT	GTGGCTTCAA	AGTGTGTGT	GTGCAGGGGG	TGGGTGTGTT	1740
AAGCTGTGGT	TAGGGAAGCG	TGTGTGCGCG	TGCTGTGGGG	CATGTGAGAT	GAGTGACTGC	1800
CGGTGAATGT	GTCCACAGTT	GAGAGGTTGG	AGCAGGATGA	GGGAATCCTG	TCAACATCAA	1860
TAATCACTTG	TGGAGCGCCA	CTTGGCCCAA	GACGCCACCT	GGCGGACAG	CAGGAGCTCT	1920
CCATGGCCAG	GGGCTCTGTG	TGCATGTTCC	CTGTCTGGTG	CCCTTTTGCC	CGCTTCTTGC	1980
AAACCTCACA	GGGTCCCTAC	ACAACAGTGC	CTCCAGGAAG	CAGCCCTCTG	GAGGAGAGGG	2040
AAGGAAATG	GGGATGGCTG	GGGCTCTCTC	CATCCTCTTT	TTCTCTTTCG	CTTGCATGG	2100
CTGGCCCTTC	CTCCAAATAC	CTCCATTTCC	CTGCTGCCAG	CCCTTTTGCC	ATAGCTGAT	2160
TTTGTGGAGG	AGGAAGGGGC	GATTTGAGGG	AGGAAGGGAG	AAAGCTTATG	GCTGGGTCTG	2220
GTCTTCTTCC	TGCCAGAGGG	GTCTTACTGT	TCCAGGGTGG	CCCCAGGGCA	GGCAGGGGCC	2280
ACACTATGCC	TGGCCCTCTG	TAAAGGTGAC	CCCTGCCATT	TACCAGCASC	CTTGGCATGT	2340
TCTTGCCTCA	CAGGAATAGA	ATGGAGGGAG	CTCCAGAAAC	TTTCCATCCC	AAAGGCAGTC	2400
TCCGTGGTGT	AAGCAGACTG	GATTTTGTCT	CTGCCCTTGA	CCCTTGTCTC	CTCTTTGAGG	2460
GAGGGGAGCT	ATGCTAGGAC	TCCAACTTCA	GGGACTCGGG	TGGCCTGCGC	TAGCTTCTTT	2520
TGATACTGAA	AACCTTTAAG	GTGGGAGGGT	GGCAAGGGAT	GTGCTTAATA	AATCAATTCC	2580
AAGCCTCAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAA			

Seq ID NO: 92 Protein sequence
Protein Accession #: Bos sequence

1	11	21	31	41	51	
MAGLAARLVL	LASAAALASG	SGQDRBPVVR	DCVLQCEEQN	CSGGALNHFR	SRQPIYMSLA	60
GWTCRDDCKY	ECMVVTVGLY	LQEGHKVPQF	HGKWPFSRFL	FFQEPASAVA	SFLNGLASLV	120
MLCRYRTFPF	ASSPMYHTCV	AFAMVSLNAN	FWSITVHTRD	TDLTERMDYF	CASTVILHSI	180
YLCCVETVGL	QHPAVVSAPR	ALLLLMLTVH	VSYLSLIRFD	XGYNLVANVA	IGLVNVVWML	240
AMCLNQRRL	PHVRKCVVVV	LLQLGLSLLS	LIDFPPLFWV	LDAHAIWHIS	TIPVHVLFPS	300
FLEDDSLYLL	KESEDKPKLD					

Seq ID NO: 93 DNA sequence
Nucleic Acid Accession #: NM_033419.1
Coding sequence: 18..980

1	11	21	31	41	51	
CGAGCCAGGG	AGAAAGGATG	GGCGGCTGG	CGGCGCGGTT	GGTCTGTCTA	GCTGGGGCAG	60
CGGCGCTGGC	GAGCGGCTCC	CAGGGCGAAC	GTGAGCCGGT	GTACCGGAC	TGCGTACTGC	120
AGTGCAGAGA	GCAGAACTGC	TCTGGGGGCG	CTCTGAATCA	CTTCCGCTCC	CGCCAGCCAA	180
TCTACATGAG	TCTAGCAGGC	TGGACCTGTC	GGGACGACTG	TAAATATGAG	TGTATGTGGG	240
TCACCGTTGG	GCCTTACCTC	CAGGAAGGTC	ACAAAGTGCC	TCACTTCCAT	GGCAAGTGGC	300
CCCTCTCCCG	GTCTCTGTTC	TTTCAAGAGC	CGGCATCGGC	CGTGGCCTCG	TTTCTCAATG	360
GGCTGGCCAG	CTGTGTGATG	CTCTGCGGCT	ACCGCACCTT	CGTGCCAGCC	TCTTCCCTCA	420
TGTACCAACAC	CTGTGTGGCC	TTCGCTGGGG	TGTCCCTCAA	TGCATGGTTC	TGGTCCACAG	480
TTTTCCACAC	CAGGGACACT	GACCTCACAG	AGAAAATGGA	CTACTTCTGT	GCTTCCACTG	540
TCATCTTACA	CTCAATCTAC	CTGTGTGCGG	TCAGGACCGT	GGGGCTGACG	ACCCAGCTG	600
TGTACAGTGC	CTTCCGGGCT	CTCTGCTGCG	TGATGTCTGAC	CGTGCACTGC	TCTTACCTGA	660
GGCTCATTCG	CTTCCGATAT	GGCTTACAAC	TGGTGGCCAA	CGTGGCTATT	GGCCTGGTCA	720
AGTGTGTGTG	GTGGCTGGCC	TGCTGCTGCT	GGAACTGACG	CGGGCTGGCT	CACGTGGGCA	780
AGTGGGTGGT	GGTGGTCTTG	CTGCTGCGGG	GGCTGTCCCT	CTCTGAGCTG	CTTGAATGCC	840
CACGCTCTCT	CTGGGTCTTG	GATGCGCCATG	CCATCTGGCA	CATCAGCACC	ATCCCTGTTC	900
ACGTCTCTCT	TTTCACTTCT	CTGGAAGATG	ACAGCCTGTA	CTTGTCTGAG	GAATCAGAGG	960
ACAAATTCAG	CTTGGACTGA	AGACCTTGGG	GCGAGTCTGC	CCCACTGGGG	ATCCCTGCCC	1020
CGCCCTGCTG	GGCTCTCTTC	TCCCTCTAAC	CTTGTGATG	ATTTTCTCTT	TTCACCTTCT	1080
TGAATCTTGA	CATGAAGGAT	GTGGGCCCGG	AATCATGTGG	CCAGCCACCC	CCCTGTGGCC	1140

CCTCACCCAGC CTGTGGAGTCT GTTCTAGGGA AGGCCCTCCCA GCATCTGGGA CTCGAGAGTG 1200
 GGCAGCCCCCT CTACCTCCTTG GAGCTGAACI GGGGTGGAAC TGAGTGTGCT CTTAGCTCTA 1260
 CCGGAGGAC AGCTGCTCTGT TTCTCCTCCA TCAGCTCCTT CCCCACATCC CCAGCTGCCT 1320
 GGCTGGGTCC TBAAGCCCTC TGTCTACCTG GGAGACCAGG GACCACAGGC CTTAGGGATA 1380
 CAGGGGGTCC CCTTCTGTGA CCACCCCTCA CCTCCTCCA GGACACCACT AGGTGGTGCT 1440
 GGATGCTTGT TCTTTGGCCA GCCAAGGTTT ACCTGATTC TCCCATGGG ATCTTGAGGG 1500
 ACCAAGCTGC TGGGATTGGG AAGGAGTTT ACCTGACCA TTGCCCTAGC CAGGTTCCCA 1560
 GGAGGCCCTCA CCATACTCCC TTTCAGGGCC AGGCTCCAG CAAGCCAGG GCAAGGATCC 1620
 TGTGTGCTG TCTGGTTGAG AGCTTGCCAC CGTGTGTGCG GAGTGTGGGC CAGGCTBAGT 1680
 GCATAGGTGA CAGGCGCTGT AGCATGGGCC TGGGTGTGTC TGAGCTCAGG CCTAGGTGCG 1740
 CAGTGTGGAG ACGGGTGTG TCGGGGAAGA GGTGTGGCTT CAAAGTGTGT GTGTGCAGGG 1800
 GGTGGGTGTG TTAGCGTGGG TTAGGGGAAC GTGTGTGCGC GTGCTGGTGG GCATGTGAGA 1860
 TGAGTGAATG CCGGTGAATG TGTCCACAGT TGAGAGGTG GAGCAGGATG AGGGAATCCT 1920
 GTCAACATCA ATAATCACTT GTGAGAGGCT AGCTCTGCC CAGGCGCCAC CTGGCGGGAC 1980
 AGCAGGAGC CAGGCGCTGT CAGGCTGCCT GTGTGCATGT TCCCTGTCTG GTGCCCCCTT 2040
 GCCCGCTCC TCAAACTTC ACAGGCTCCC CACACAACAG TGCCCTCCAG AAGCAGCCCC 2100
 TCGGAGGCG AGGAAGGAAA ATGGGATGG CTGGGGCTCT CTCCATCCTT CTTTCTCCT 2160
 TGCTTGGCA TGGCTGGCTT TCCCTCCAA AACCTCCAT CCCCCTGCTG CAGCCCCCTT 2220
 GCCATAGCTT GATTTTGGGG AGGAGGAAGG GCGGATTGA GGGAGAGGG GAGAAAGCTT 2280
 ATGGCTGGG CTGTTTCTT CCTTCCAG AGGCTCTTAC TGTTCAGGG TGGCCCCAGG 2340
 GCAGGAGGG GCCACATAT GCCTGGGCC TGGTAAGGT GACCCCTGCC ATTTACCAGC 2400
 AGCCCTGGCA TGTCTCTGCC CCACAGGAAT AGAATGGAG GAGCTCCAGA AACTTCCAT 2460
 CCCAAAGGCA GTCTCGTGG TTGAAGCAGA CTGGATTTT GCTCTGCCCT TGACCCCTTG 2520
 TCCCTCTTT AGGGAGGGGA GCTATGCTAG GACTCCAAAC TCAGGAGCTC GGGTGGCTG 2580
 CGCTAGCTTC TTGTATACT GAAACTTTT AAGTGGGAG GGTGGCAAGG GATGTGCTTA 2640
 ATAAATCAAT TCCAGCTCT AAAAAAAAAA AAAAAAAAAA AA

Seq ID NO: 94 Protein sequence
 Protein Accession #: NP_219487.1

1 11 21 31 41 51
 MKDVPSPSCG QPTPCWPSPA LESVLGKASQ HLGLESQGQL YLLELNWGGT ECALSTGRT 60
 AACFLPISLL PTSPAALWGP EALCLPGRPG TTGLRDTGGP LLLPPPTLLQ DTRWCMMLV 120
 LWPAKVHSDS PHGILRDQAA GIGKEFRPDH CPSQVPRRPH HTFFQGGSS KPRARILCCC 180
 LVESLPFCVG SVQQAECIGD RAVSMGLGVC ELRPRCAVWR RVLGKRCGF KVCVCRGWVC

Seq ID NO: 95 DNA sequence
 Nucleic Acid Accession #: XM_090469
 Coding sequence:

1 11 21 31 41 51
 ATGGGGTTTG GAGACCAGGG AACGGTGGAA GGGAGCCTAG GAACGTGAA AGATAGCTCC 60
 TGCCCGCTGC TGGACAGGAC GCAGGTGAAC AAGCCCTGCC TGCTGGAGGC TGCCAGCAC 120
 AGCCACACAC CCTGTGATGG ACAGTGGTTC AACAGCCACC CATCAGGTGC TCATCAGAAA 180
 CCACCTGAGG TGAATAATGTT TGGAGCCAGT CAAGGTTTGC TGACAATGGA AACAAACCA 240
 TCCCTGGCAC AAGGCACAGG CTGCTCAGTG TGCTGGGAGG TGAACGGCCC AGGGTTGGAA 300
 ATGTCTCTTC CTGGAGTGCT CAGTGTCTGAT GCTGGCCAGG TGGAGCACAG AAGACAAATG 360
 AACTCGCAG ATACTGACTG GGGAACTGAG GAGGGACCAG GAAGTCTAGA TTTCCGCGTG 420
 GCTGCCACCC AGGAAGACCA TTCTTTTTC AAAGACATCA AGCACACAAG TACGTTCCAG 480
 CAATCAGTTC AGCAGCAGAA TTGCATTTAC AGCCCCAGAG AAAAAACCTG TGGGAATGTC 540
 AAGGCACCTT GGGGCCACCC AAGGCGAGAG GCCCCGCTTG CCCCAGCAG GGCCTGGCGG 600
 CTCAGCGGCA GCCCGGGGCC GTCCCAACGT ACACCCATGG CGCCCTTCCC GACCTCTGAC 660
 CGAGAGCTAG ACGGCCAGAG CCGCCGCCCC GGGCTGGGGA GCTCGGCAGC AGCCCTTCAC 720
 TGCCCTCCCG CGGTCCAGA AGCCAGAAA CTCTCTTCC TGCTGCCCTT TTATCCAGAT 780
 GGCAGGCCAC CACTAAGGA CATCTGACG AGCTCCAGC ACAAGGCACC TGGGAAGAGA 840
 CTGCATCAAG AATCACCAGG TTCATTTCCA ATAGGTTTAT GTAAAGAAC AGCTGAAACA 900
 TGATCAAGG GGCCTCTGGG TGTGGCTGCT GAAGCAGCAT TTAATTTACA GTTTTCCAGT 960
 GAATGA

Seq ID NO: 96 Protein sequence
 Protein Accession #: XP_090469

1 11 21 31 41 51
 MGFQDQSTVE GSLGTSKDSB CRVLDRQVN KPCLEAAKH SPHPLEGGQSV NSEPSGAHQK 60
 FFEVKMFQAS QQLLTMETNQ SLAQGTGCSV CNEVNGPGLB MSLPGVLBAD AGQVEHRRQM 120
 NSADTDNGBE BPGSLDFAV AAHQEDTFFL KDIKHTSTFR QSVQQNCIY SPREKPCGNV 180
 RAPCAPPRRE APLALSRPWR LSRSPAPSPR TPMAPFTSD RELDAFGPPP GLRSSAAAPH 240
 CLPAAPPAQK LFFLLFLYED GSPPPKDILQ TLQHKAPGKR LHQSSPQSFY IGLCNRTARP 300
 CIRGALGVAA EAAPFLQFSS E

Seq ID NO: 97 DNA sequence
 Nucleic Acid Accession #: NM_003474.2
 Coding sequence: 37..3036

1 11 21 31 41 51
 CACTAACGCT CTCTCTAGTC CCGGGGCCAA CTCGGACAGT TTGCTCATTT ATTGCAACGG 60
 TCAAGGCTGG CTGTGCTCCAG AAGGCGCGCG GCGGAGCGCA CGCACACACA CCGGGGGGAA 120
 CTTTITTAATA AATGAAGGCT TAGAAGAGCT CAGCGCGCGC GCGGGCGGTG CGGAGGGGCT 180
 CCGGAGCTGA CTCGCGGAGG CAGGAATCC CTCGCTGCGC GACGCGCGGC CCGCTCGGCG 240
 GCGGCGGTGG GATGGTGCAG GCCTCGCGCG CCGGCGCGAG AGCTGTCTGA CTGAAGGCGG 300

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

```

GCGACGATGG CAGCGCBCCC GCTGCCCGTG TCCCCCGCCC GCGCCCTCCT GCTCGCCCTG 360
GCCGGTGTCT TGTCTGCGCC CTGCGAGGCC CBAGGGGTGA GCTTATGGAA CGAAGGAAGA 420
GCTGATGAAG TTGTCACTGC CTCTGTTCCG AGTGGGGACC TCTGGATCCC AGTGAAGAGC 480
TTGCACTCCA AGAATCAATCC AGAAGTGTCTG AATATTGAC TACAACGGGA AAGCAAGAGAA 540
CTGATCATAA ATCTGGAAAG AAATGAAGGT CTCTTGCCA GCAGTTTCAC GGAAACCCAC 600
TATCTGCAAG ACGGTACTGA TGTCTCCCTC GCTCGAAATT ACACGGTAAT TCTGGGTAC 660
TGTTACTACC ATGGACATGT ACGGGGATAT TCTGATTGAG CAGTCAGTCT CAGCACGTGT 720
TCTGGTCTCA GGGGACTTAT TGTGTTTGA AATGAAGCT ATGTCTTAGA ACCAATGAAA 780
AGTGCACACA ACAGATACAA ACTCTTCCCA GCGAAGAGC TGAAGAAGGT CCGGGGATCA 840
TGTGGATCAC ATCTCAACAC ACCAAACCTC GCTGCAAGA ATGTGTTTCC ACCACCTCT 900
CAGACATGGG CAGAAGGCA TAAAGAGAG ACCCTCAAG CAATCAAGTA TGTGGAGCTG 960
GTGATCGTGG CAGACAACCG AGAGTTTCAG AGGCAAGGAA AAGATCTGGA AAAAGTTAAG 1020
CAGCGATTAA TAGAGATTGC TAATCACGTT GACAAGTTT ACAGACCCTT GAACATTGG 1080
ATGTGTGTGG TAGCGGTGGA AGTGTGGAAT GACATGGACA AATGCTCTGT AAGTCAGGAC 1140
CCATTACCCA CTCTCACTGA ATTCTGGAC TGGAGGAAGA TGAAGCTTCT ACCTCGCAA 1200
TCCCATGACA ATGCGCAGCT TGTCACTGGG GTTTATTTC AAGGCAACAC CATCGGCATG 1260
GCCCCAATCA TGAGCATGTG CACGGCAGAC CAGTCTGGGG GAATTGICAT GGACATTCA 1320
GACAATCCCC TTGTTGCAAG CGTGACCTTG GCACATGAGC TGGGCCACAA TTTCGGGATG 1380
AATCATGACA CACTGGACAG GGGCTGTAGC TGTCAAAATGG CGTTTGAGAA AGGAGGCTGC 1440
ATCATGAACG CTTCCACCGG GTACCCATT CCATGCGTGT TACAGAGTTG CAGCAGGAAG 1500
GACTTGGAGA CCAGCCTTGA GAAAGGAATG GGGTGTGCTC TGTTAACCT GCGGGAAGTC 1560
AGGAGTCTTT TCGGGGCCCA GAAGTGTGGG AACAGATTG TGAAGAAGG AGAGAGTGT 1620
GACTGTGGGG AGCCAGAGGA ATGTATGAAT CGCTGCTGCA ATGCCACCAC CTGTACCTTG 1680
AAGCCGAGCG CTGTGTGCGC ACATGGGCTG TGTGTGAAG ACTGCCAGCT GAAGCCTGCA 1740
GGAACAGCGT CAGGCGCTC CAGCAACTCC TGTGACCTCC CAGAGTTCTG CACAGGGGCC 1800
AGCCCTCACT GCCCAGCCAA CGTGTACTTG CACGATGGGC ACTCATGTCA GGATGTGGAC 1860
GGCTACTGCT ACAATGGCAT CTGCCAGACT CACGAGCAGC AGTGTGTAC ACTCTGGGA 1920
CCAGGTGCTA AACCTGCCCC TGGGATCTGC TTTGAGAGAG TCAATTCTGC AGGTGATCCT 1980
TATGGCAACT GTGGCAAAAT CTGGAAGAGT TCCTTTGCCA AATGCGAGAT GAGAGATGCT 2040
AAATGTGGAA AAATCTCATG TCAAGGAGGT GCCAGCGGCG CAGTCATTGG TACCAATGCC 2100
GTTTCCATAG AAACAACAT CCCTCTGAGC CAAGGAGGCC GGATTCGTG CCGGGGGACC 2160
CAGCTGTACT TGGCGATGA CATGCCGAC CAGGGCTTG TGTGTGAGG CACAAGTGT 2220
GCAGATGGAA AAATCTGCTT GAATCGTCAA TGTCAAATA TTAGTGTCTT TGGGTTTAC 2280
GAGTGTGCAA TGCAGTGCAC CGGCAGAGGG GTGTGCAACA ACAGGAGGAA CTGCCACTGC 2340
GAGGCCACT GGCACCTCC CTCTGTGAC AAGTTTGGCT TTGGAGGAAG CACAGACAGC 2400
GGCCCCATCC GGCAGCAGA TAACCAAGGT TTAACCATAG GAATTCGGT GACCATCTG 2460
TGCTCTCTG CTGCCGATT TGTGGTTTAT CTCAAAAGGA AGACCTGAT ACAGTCTG 2520
TTTACAAATA AGAAGACCAC CATTGAAAAA CTAAAGTGTG TGCGCCCTTC CCGGCCACCC 2580
CGTGGCTTCC AACCTGTGCA GGCTCACTTC GGCCACCTTG GAAAGGCCCT GATGAGGAAG 2640
CGCCAGAT CTATACCCAC GAAGGACAAT OCCAGGAGAT TGCTGAGTG TCAGAAATGT 2700
GACATCAGA GACCCCTCAA CGGCTGAAAT GTCCCTCAGC OCCAGTCAAC TCAGCGAGTG 2760
CTTCTCTCCC TCCACCGGCG CCAAGTGCA CCTAGGCTCC CTGCCAGAC CCGCCAGCC 2820
AAGCCTGCAC TTAGGCAGGC CCGGGGACC TGTAAGCCAA ACCCCCTCA GAAGCCTCTG 2880
CTTGACAGAT CTCTGGCCAG AACAACTCGG CTCACTCATG CCTTGGCCAG GACCCAGGA 2940
CAATGGGAGA CTTAGCTCCG CCTGGCACCC CTCAGACCTG CTCACAAATA TCCACACAA 3000
GTGCCAGAT CCACCCACAC CGCTTATATT AAGTGAGAAG CCGACACCTT TTTTCAACAG 3060
TGAAGACGA AGTTTGCAT ATCTTTCAGC TCCAGTTGGA GTTTTGTGA CCAACTTTTA 3120
GGATTTTTTT TAATGTTTAA AACATCATTA CTATAAGAAC TTTGAGCTAC TGCCTGCTAGT 3180
GCTGTGCTGT GCTATGCTGC TCTGTCTACT TGCAACGGTA CTGTAAATTT ATTAATTTAT 3240
GCAGAAATGT GATTACAGTG CAGTGGCTG TAGTAGGCAT TTTTACCTAT ACTGAGTTT 3300
CCATGGCAGG AAGGCTGTGT GTGCTTTTAG TATTTTAGTG AACTTGAAT ATCTGTCTG 3360
ATGGGAATCT GCACAGGATG TGTGCTTT CTGATCAAGG CCTTATGGA AAGCAGTCCC 3420
CCAATACCC CAGCTGTGTC TTATGGTACC AGATGCAGCT CAAGAGATCC CAAGTAGAAT 3480
CTCAGTTGAT TTCTGGATT CCCCATCTCA GGCAGAGGCC AAGGGGCTTC AGGTCCAGG 3540
TGTGTTTGGC TTTGAGGAG GCGCTGTGCC CCTTGACAAC TGGCAGGCG GCTCCAGGG 3600
ACACCTGGGA GAAATCTGCG TCTGGGCCAG GAAGCTTGG TGAGAACCTG GGTTCAGAC 3660
AGGAATCTTA AGGTGTAGCC ACACAGGAT AGAGACTGGA AACTAGACA AGCCAGAACT 3720
TGACCTGAG CTGACCAGGC GTGAGCATGT TTGGAAGGG TCTGTAGTGT CACTCAAGGC 3780
GGTGCTTGAT ABAAATGCCA AGCACTTCTT TTTCTGCTG TCCTTTCTAG AGCACTGCCA 3840
CCAGTAGGT ATTAGCTTG GGAAGGTGG TGTTCGTGA AGAAAOCTAC TGCCAGGCA 3900
CTGCAACCG CCACCTCCCT ATACTGCTG GAGCTGAGCA ATCAACACA AACTGTAATA 3960
CAATGATCTT GTATTGAGC AGATGAGGAC TTTCCATGG ACCACAATA TTTTCAGATG 4020
TGAACCAATA ACCAGATCTA GTCAATCAG TCTGTTTACT GCAAGGTICA ACTTATTAA 4080
AATTAGGCAG ACTCTTTATG CTGCAAAAA CTACAACCAA TGAATGTGA TGTTCATGG 4140
TATAGTTTAT GTCTGCTATC ATTATTGTA GATATTGGAC AAAGAACCTT CTCTATGGG 4200
CATCTCTT TTCCAACTTG GCTGAGGAA TCTTTAAAG ATGCTTTTAA CAGAGTCTGA 4260
ACCTATTCTT TAAACCTTG CAACCTACCT GTTGAGCATC ACAGAATGTG ATAAGGAAAT 4320
CACTTGTCTT ATCACTTCC TAAATATTAT GAGATGTGG TTTGGGAGCA TCCCCTTGA 4380
CTCTTCACTC TTAATATGCC TGACTAGGA GGCATGTTTC ACAGGCTCTT TAAAGTACT 4440
AATGGCATGA GAAATACAAA AATACTCAGA TAAGGTAAAA TGCCATGATG CCTCTGCTT 4500
CTGGAGTGGT TTTCACTTA GAAGACAATT GACAACAGTT ACATAATTCA CTCTGAGTGT 4560
TTTATGAGAA AGCCTCTCTT TGGGTCAAC AGTTTTCCTA TGCTTTGAAA CAGAAAAATA 4620
TGTACCAAGA ATCTGTGTTT GCCTTCCGGA AAACAAACT GCATTTCAT TTCCCGTGT 4680
TCCCACATGT ATCTAGGCAA CATAGTATTC ATGACTATGG ATAAACTAAG CACGTGACAC 4740
AAACACACAC AAAAGGGAAC CCAGCTCTAA TACATTCCAA CTCGTATAGC ATGCATCTG 4800
TTATTCTATA GTTATTAGT TCTTTAAAT GTAAAGCCAT GCTGGAAAT AATACGTCTG 4860
AGATACATAC GAATTTACG TAATGATTA CACTTGTGTA TTGTACTAAA GCCAAACATA 4920
TATATACTAT TAAAAAGGT TACAGATTT TATGTTGAT TACGTGGGCA TTGTCTTTT 4980
AGATGCCCAA ATCTTAGAT CTGGCATGTT AGCCCTCTCT CCAATTATTA GAGGATATGA 5040
ACCAAAAAA AAAAAAAA AA
  
```

Seq ID NO: 98 Protein sequence
 Protein Accession #: NP_003465

1 11 21 31 41 51
 1 MAARPLPVSP ARALLHLAG ALLAPCEARG VSLWNEGRAD EVVSASVRSR DLWIFVKSF 60
 5 SKNHPEVLNI RIQRESKELI INLERNEGLT ASSFTETHYL QDGTDVSLAR NYTVILGHY 120
 YHGHVRGYSD SAVSLSTCSG LRGLIVFENE SYVLEPMKSA TNRYKLFPK KLSVREGSG 160
 SHHNTFNLAA KNVFPSPST WARRHKRETL KATKYVELVI VADNREFQRQ GKOLEKVKQR 240
 LIETANHVDK FYRPLNIRIV LVGVEVWDM DXCSVSQDPF TSLHEFLDWR KMKLLPRKSH 300
 DNAQLVSGVY FQGTIGMAP IMSMCTADQS GGIVMDHSDN PLGAAVTLAH ELGHNFGMNH 360
 10 DTLDRCCSCQ MAVKGGCIMI NASTGYFFPM VFSSCSRKDL ETSLEKMGV CLFNLEFVRE 420
 SFGGQKCGNR FVEBEGCDC GRPERCMNRC CNATTCTLKP DAVCAGLCC EDCQLKPA 480
 ACRDSSNCD LPEPCTGASP HCPANVYLHD GHSCQDQDGY CYNGICQTHE QQCCTLWGP 540
 AKPAEGICFE RVNSAGDPYG NCGKVSXSF AKCBMRDAKC GKIQCQGGAS RPIVGTNAV 600
 IETNIPLOQ GRILCRGTHV YLGDDMPDPG LVLAGTKCAD GKICLNRCQ NISVFGVRE 660
 15 AMQCHGRGVC NNRKNCHCEA HWAPPFCDF GPSSSTDSGP IRQADNQGLT IGIIATITL 720
 LAAGFVYLK RKTILIRLLFT NKKTITIEKLR CVRPSRPPRG FQPCQHLGH LGKGLMRKP 780
 DSYPEKDNER RLQCCQNVDI SRPLNGLNVP QPQSTQVLP FLHRAPRAPS VPARPLPAK 840
 ALRQAGGTC ENPQKPLPA DPLARTTRLT HALARTPGQW ETGLRLAPLR PAFQYFHV 900
 RSTHTAYIK

20 Seq ID NO: 99 DNA sequence
 Nucleic Acid Accession #: NM_003714
 Coding sequence: 135..1043

25 1 11 21 31 41 51
 1 GAGGAGGAGG GAAAGGCGA GCAAAAGGA AGAGTGGAG GAGGAGGGA AGCGCGAAG 60
 GAGGAGGAGG AGGAGGAGGA AGAGGGGAGC ACAAGGATC CAGGTCTCCC GACGGGAGGT 120
 TAATACCAAG AACCATGTGT GCGGAGCGGC TGGGCCAGT CATGACCTG GCTTTGGTGT 180
 30 TGGCCACCTT TGACCGGCG CGGGGAGCCG ACBCCACCA CCCACCCGAG GGTCCCAAG 240
 ACAGGAGCTC CACGAGAAA GCGCGCTGT CCTGCAGAA TACAGCGAG ATCCAGCACT 300
 GTTTGGTCAA CGCTGGCGAT GTGGGTTGT GCGTGTGTA ATGTTTCGAG AACCACTCT 360
 GTGAGATTG GCGCTTACAT GGGATTGCA TGACTTTCT GCACAAAGCT GGAATAATTG 420
 ATGCCAGGG CAAGTCATT ATCAAGACG CCTGAAATG TAAGGCCAC GCTCTCGGC 480
 35 ACAGGTTCCG CTCCTAAGC CGGAAGTGC CGCCATCAG GGAATGGTG TCCAGTTGC 540
 AGCGGGAATG CTACCTCAG CACGACCTGT GCGCGCTGC CCAGGAGAAC ACCCGGGTGA 600
 TAGTGGAGAT GATCATTTC AAGGACTTGC TGCTGCAGA ACCCTACGT GACCTCGTGA 660
 ACTTCTGCT GACCTTGGG GAGGAGGTGA AGGAGGCCAT CACCCACAGC GTCCAGGTT 720
 AGTGTGAGCA GAATGGGGA AGCTGTGCT CATCTTAG CTTCGACC TCGCCATCC 780
 40 AGAAGCTCC CACGCGCCC CCGAGCGCC AGCCCGAGT GGACAGAAC AAGCTCTCCA 840
 GCGCCCAACA CCGGGAAGCA GGACATCAC TCCAGAGCC CAGCAGTAG GAGACTGGCC 900
 GAGGTGCCAA GGGTGAAGCA GGTAGCAAGA GCCACCAA CCGCCATGCC CGAGGCAGAG 960
 TCGGGGCGCT TGGGCTCAG GACCTTCCG GAAGCAGCA GTGGGAAGAC GAACAGTCTG 1020
 AGTATTCTGA TATCCGAGG TGAATGAAA GGCCTGCCA CGAATCTTT CCTCCAGGCC 1080
 45 GTCCATTTC TTATCTATG ACATTCCAA ACATTACCA TTAGAGAGG GGGATGTCAC 1140
 ACGCAGGAT CTGTGGGAC TGTGGACTTC ATCAGGTGT GTGTTCGCG AACGGACAGG 1200
 TGAGATGGAG ACCCTTGGG CGGTGGGTC TCAGGGTGC CTGGTGAAT CTGCACTTAC 1260
 ACCTACTCAA GGGAGCGGC CGCGTTATC CTGTACCTT TGTCTCTCT CCATCTGTG 1320
 AGTCAGTGG TGTGCGCGC TCTGTGTGG GGGAGGTGA CCAGGGAGG GCAGGGCAAG 1380
 50 CAGGGCGCCC CAGAGCTGG CCACACAGT GGTGCTGGC CTCGCCCGA AGCTTCTGGT 1440
 GCAGCAGCT CTGTGCTGT CTCGCGGAA GTCAGGGCG CTGGATTCCA GACAGGAGT 1500
 GAATGTAAA ATAAATATC CTAGAGATG AGGAGAAGG TGGAGAGGAG CCAGGGCGCG 1560
 AGCGGGTGT TGGTGCCTA CTGAAATTCA GTTCTTGTG TGGGCGCTG CGGTTCAAG 1620
 CTCTGTGCA GGGTGGAGG AGGAGTGA TTTCTATGT TAATTCTGA GCCATTGTAC 1680
 55 TGTCTGGCT GGGGGGACA CTGTCCAAG GAGTGGCCC TATGAGTTA TATTTAAC 1740
 ACTGCTTCAA ATCTCGATT CACTTTTTT ATTTATCCG TTATATCTAC ATATCTGTCA 1800
 TCTAATAAAA TGGCTTTCAA ACAAGCAAC TGGGTCTTA AAACAGCTC AAAGGGGGT 1860
 TAAAAAATA AAAACAGCC CATCTTTGA GCGTGAATT TCTTTTTTT AAGTTCTATT 1920
 TTAAGAGCTA TCAACAGCG ACATAGCCAT ACATCTGACT GCTGACATG GACTCCGSC 1980
 60 CACTTGGGG AAACCTTATA CCCAGAGGA AATACACAC TGGGGAGTAC ATTTGACAAA 2040
 TTTCCTTAG GATTTCTTA TCTCACCTTG ACCCTCAGC AAGATTGGTA AAGCTGGTTC 2100
 CTGCGGATTC CAGGAGACCT AGCTGGAAC CTGGCTTCT CATGTGAGG GATGGGAAG 2160
 GAAAGAAGAG AATGAAGACT ACTTAGTAAT TCCCATCAG AATGCTGAC CTTTACATA 2220
 AAATCAAGGA GACTGCTGAA AATCTCTAG GACACAGGATT TTCCAGATCC TAATTGAAA 2280
 65 TTTAGCAATA AGGAGAGGAG TCCAAGGGGA CAAATAAAG CAGAGAGAGA GAGAGAGAA 2340
 GGGAGAGGAA GAAAGAGAG AGAGAAAAGA GCCTCGTCCC

Seq ID NO: 100 Protein sequence
 Protein Accession #: NP_003705

70 1 11 21 31 41 51
 1 MCAERLQDEM TLALVLATFD PARGTDATNP PEGEQDRSSQ QKGRSLQNT AEIQHCLVNA 60
 GDVCGGVFEC FENNSCEIRG LKHICMTFLH NAGKFDAGK SPIKDALCK ABALRHFRFC 120
 75 ISRKPAIRE MVSQLRECY LKHDLCAAQ ENTREVIVEMI HFKDLLLHEP VVDLVNLLLT 180
 CGEEVKEAIT HSVQVQCEBN NGSLCSILSF CTSATQKPPT APPERQPOVD RTKLSRAHAG 240
 EAGHHLPEPS SRETGRGAGK ERGSKSHENA HARGEVGGLG AQGPSGSSSW EDEQSEYSDI 300
 RR

80 Seq ID NO: 101 DNA sequence
 Nucleic Acid Accession #: NM_005940
 Coding sequence: 23..1489

1 11 21 31 41 51

5
 10
 15
 20
 25
 30
 35
 40
 AAGCCACGCA GCCCCGGGGC GGATGGCTCC GGCOCCTGG CTCGSCAGCG CGGCCGCGGG 60
 CGCCCTCTCT CCCCCGATGC TGCTGCTGCT GCTCCAGCCG CCGCCGCTGC TGGCCCGGGC 120
 TCTGCGCCCG GACGCTCCACC ACCTCCATGC CGAGAGGAGG GGGCCACAGC CCTGGCATGC 180
 AGCCCTGGCC AGTAGCCCGG CACCTGCCCC TGCCACGCG AGAGCCCCCG GGCCCTGCCG 240
 CAGCCTCAGG CTTCCCGCGT GTGGCGTGCC CGACCCATCT GATGGGCTCA GTGCCCGCAA 300
 CCGACAGAAG AGGTTCTGTC TTCTGCGCGG GCGCTGGGAG AAGACGGACC TCACCTACAG 360
 GATCCTTCGG TTCCCATGGC AGTTGGTGCA GGAGCAGGTG CGGCAGACGA TGGCAGAGGC 420
 CCTAAAGSTA TGGAGCGATG TGACGCCACT CACCTTTACT GAGGTGCACG AGGGCCGTGC 480
 TGACATCATG ATCGACTTCG CCAGGTAATG GCATGGGGAC GACCTGCCGT TTGATGGGCC 540
 TGGGGGCGTC CTGGCCCATG CCTTCTTCCC CAAGACTCAC CGAGAAGGGG ATGTCCACTT 600
 CGACTATGAT GAGACCTTGA CTATCGGGGA TGACCAAGGC ACAGACTTGC TGCAGGTGGC 660
 AGCCCATGAA TTTGGGCACG TGCTGGGGCT GCAGCACACA ACAGCAGCCA AGGCCCTGAT 720
 GTCCGCTTTC TACACCTTTC GCTACCCACT GAGTCTCAGC CCAGATGACT GCAGGGCGGT 780
 TCAACACCTA TATGSCCAGC OCTGGCCAC TGTCACTCC AGGACCCAG CCTGGGCC 840
 CCAGGCTGGG ATAGACACCA ATGAGATTGC ACCGCTGGAG CCAGACGCCC CGCCAGATGC 900
 CTGTGAGGCC TCCTTTGAAG CGGTCTCCAC CATCGAGGC GAGCTCTTTT TCTTCAAAGC 960
 GGGCTTTGTC TGGCGCTTCC GTGGGGGCCA GCTGCAGCCC GGCTACCCAG CATTGGCCTC 1020
 TCGCCACTGG CAGGGACTGC CCAGCCCTGT GGAAGCTGCC TTCAGGATG CCGAGGGCCA 1080
 CATTGGTTC TTCCAAGGTG CTCAGTACTG GGTGTACGAC GGTGAAAAGC CAGTCTGGG 1140
 CCCCACACCC CTCACCGAGC TGGGCTGGT GAGGTTCCCG GTCCATGCTG CTTTGGTCTG 1200
 GGGTCCCGAG AAGAACAGA TCTACTTCTT CCGAGGCAGG GACTACTTGC GTTCCACCC 1260
 CAGCACCCGG CGTGGCCACG GTCCCGTGCC CCGCAGGGCC ACTGACTGGA GAGGGGTGCC 1320
 CTCTGAGATC GACGCTGCC TCCAGGATGC TGATGGCTAT GCCTACTTCC TGCCTGGCCG 1380
 CCTCTACTGG AAGTTTGCCT CTGTGAAGGT GAAGGCTCTG GAAGGCTTCC CCGCTCTGCT 1440
 GGGTCTGAC TCTTTTGGCT GTGCCGAGCC TGCCAAACTT TTCTCTGAC CATGGCTTGG 1500
 ATGCCCTCAG GGGTCTGAC CCGTCCAGG CCAAGAAATAT CAGGCTAGAG ACCCATGGCC 1560
 ATCTTTGGG CTGTGGGCAC CAGGCAATGG ACTGAGCCCA TGTCTCTGTC AGGGGGATGG 1620
 GGTGGGGTAC AACCACTATG ACAAATGCGG GAGGGGCCAC GCAGGTCTGT GTCCACTGCC 1680
 AGCGACTGTC TCAGACTGGG CAGGGAGGCT TTGGCATGAC TTAAGAGGAA GGCAGTCTT 1740
 GGGACCGCT ATGCAGGTTC TGGCAAACTT GGCTGCCCTG TCTCATCCCT GTCCCTCAGG 1800
 GTAGCAACAT GGCAGGACTG GGGAACTGG AGTGTCTTTC CTGTATCCCT GTTGTGAGGT 1860
 TCCTTCCAGG GCTGVCACCT GAAGCAAGGG TGCTGGGGCC CCATGGCCTT CAGCCCTGGC 1920
 TGAGCAACTG GGCCTAGGG CAGGGCCACT TCCGAGGTTC AGGTCTTGGT AGTGCCTGTC 1980
 ATCTGTCTGC CTTCGGCTG ACAATCCTGG AAATCTGTTT TCCAGAAATC AGGCCAAAA 2040
 GTTCACAGTC AAATGGGGAG GGGTATTCTT CATGCAGGAG ACCCCAGGCC CTGAGGCTG 2100
 CAACATACCT CAATCCTGTC CCAGGCCCGA TCTCTCTGAA GCCCTTTTTC CAGCACTGCT 2160
 ATCTTCCAAA GCCATTGTAA ATGTGTGTAC AGTGTGTATA AACCTCTTC TTCTTTTTTT 2220
 TTTTAAACT GAGGATTGTC ATTAAACACA GTTGTTTTCT

Seq ID NO: 102 Protein sequence

Protein Accession #: NP_005931

45
 50
 55
 1
 11
 21
 31
 41
 51
 MAPAANLRS AARALLPPL LLLQPPPLL ARALPPDVH LHAERBPQP WHAALPSSPA 60
 PAPATQEAR PASSRPPRC GVDFPDSGLS ARNRQKRFVL SGRWRKTDL TYRILRFPWQ 120
 LVQGVQRVIM AALALVNSDV TPLTFTEVHE GRADIMLDPA RYWHGDDLPF DGPGGILAH 180
 FFPKTHREDG VRFVDWETMT IGDQDQTDLL OVAHEFGHV LGLQHTTAAL ALMSAFYTF 240
 YPLSLSDDDC RGVQHYLQGP WPTVTSRTFA LGPQAGIDTN RIAPLEPDAP PDACEAFDA 300
 VSTIRGELFF FRAGFVWRLR GQQLQPGYPA LASRWQGLP SPVDAAPEDA QGHINFPQGA 360
 QYNVYDGEKP VLGPAFLTEL GLVRFVHDA LVNGPKNKI YPFRGRDYWR FHPSTRVDS 420
 FVPRRAZDMR GVPSEIDAAF QDADGYAYFL BGRLYWKFDP VKVKALEGFP RLVPDFDFGC 480
 AEPANTFL

Seq ID NO: 103 DNA sequence

Nucleic Acid Accession #: NM_033151.2

Coding sequence: 351..4499

60
 65
 70
 75
 80
 1
 11
 21
 31
 41
 51
 ACTGGGATAA AGCAAGAAGA CTGATTTTAT GAGCAGGGGT TTGATACATC AAAGGAGATT 60
 GCCCAGGATC AAGGGTGGCG TGTGGGGGT GGGTTGGGGA GGGTGGTAG AGAAGGTTTC 120
 ACTAAGTGAT TTTGGGCTGA GGCCTGAGAA GATGTTTAAA AAGAGGGATC AAGCACAGGC 180
 TAAGGAGAGG AAGAGCAGG CACCCAAACC TCTGCATGGC CCAATATGC TCCCTGCAGG 240
 GTAGTGCCCC CTCTCTGGC TGCTCAAGGC GAGATCTAAG CTCTCTTAA CTCTGCTGT 300
 CTTTCTATAT TCTCTGATTC TGGGAAACGA AGAATTGGCA GGAACGTGAA ATGACTAGGA 360
 AGAGGACATA CTGGGTCGCC AACTCTTCTG GTGGCCTCGT GAATCGTGGC ATCGACATAG 420
 GCGATGACAT GGTTCAGGA CTTATTTATA AAACCTATAC TCTCCAGAT GSCCCTGGA 480
 GTCAGCAAGA GAGAAATCCT GAGGCTCCAG GGAGGCGAGC TGTCCACCG TGGGGAGT 540
 ATGATGCTGC CTTGAGAACC ATGATTCCTT TCCGTCCCAA GCGAGGTTT CCTGCCCCCC 600
 AGCCCTTGGG CAATGCTGGC CTGTCTCTCT ACCTCACCTT GTCATGGCTC ACGGCTTCA 660
 TGATCCAAAG CTACCGAGT CGCTTAGATG AGAACACCAT CCTCCACTG TCACTCCATG 720
 ATGCTCTAGA CAAAGATGTC CAAAGGCTTC ACCGCTTTC GGAAGAGAA GTCTCAAGGC 780
 GAGGATGGA AAGAGCTTCA GTGCTTCTG TGATGCTGAG GTTCTAGAGA ACAAGGTTGA 840
 TTTTGGATGC ACTTCTGGGC ATCTGCTTCT GCATTGCCAG TGTACTCGGG CCAATATTGA 900
 TTATACCAAA GATCCTGGAA TATTCAAGAG AGCAGTTGGG GAATGTTGTC CATGGAGTGG 960
 GACTCTGCTT TGCCCTTTT CTCTCCGAA GTGTGAAGTC TCTGAGTTTC TCTCCAGTT 1020
 GGATCATCAA CCAACGCACA GCCATCAGGT TCGAGCAGC TGTTCCTCC TTGCTCTTTC 1080
 AGAGCTCAT CCAATTTAAG TCTGTAATAC ACATCACTTC AGGAGAGGCC ATCAGCTTCT 1140
 TCAACGGTGA TGTAAACTAC CTGTTTGAAG GGGTGTGCTA TGGACCCCTA GTACTGATCA 1200
 OCTGCGCATC GCTGGTCATC TGACGATTT CTCTCTACTT CATTATTGGA TACACTGCTA 1260
 TTATTGCCAT CTATGCTAT CTCTGGGTTT TCCCACTGGC GGTATTTCAT ACAAGAAATG 1320
 CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCAACA GGCATCCCT GTGACCAATG 1380
 AAGTTCTCAC TTGCATTAG CTGATTAATA TGTACACATG GGAGAAACCA TTGCAAAAA 1440

TCATTGAAGA CCTAAGAAGG AAGGAAAGGA AACTATTGGA GAAGTGC0GG CTTGTCCAGA 1500
 GCGTACAAAG TATAACCTTG TTCAATCATCC CCACAGTGGC CACAGCGGTC TGGGTTCTCA 1560
 TCACACATCT CTTAAAGCTG AAACCTACAG CGTCAATGGC CTTACAGCATG CTGGCCTCTCT 1620
 TGAATCTCCT TCGGCTGTCA GTGTTCCTTG TGCTATTGCT AGTCAAAGGT CTCACGAATT 1680
 CCAAGTCTGC AGTGATGAGG TTCAAGAAAT TTTCTCTCCA GGAGAGCCCT GTTTTCTATG 1740
 TCAGACATT ACAAGACCCC AGCAAGCTC TGCTCTTTGA GGAGGCCACC TTGTTCATGGC 1800
 AACAGACCTG TCCCGGGATC GTCAATGGGG CACTGGAGCT GGAGAGGAAC GGGCATGCTT 1860
 CTGAGGGGAT GACCAGGCGT AGAGATGCCC TCGGGCCAGA GGAAGAAGGG AACAGCCTGG 1920
 GCCCAGAGTT GCACAAGATC AACCTGGTGG TGTCCAAGGG GATGATGTTA GGGGTCTGGC 1980
 GCACACGAGG GATGGTGAAG AGCAGCCTGT TGTCAGCCAT CCGGAGGAG ATGCACCTGC 2040
 TCAGGGGCTC GGTGGGGGTG CAGGGAAGCC TGCCCTATGT CCCCAGCAG GCCTGGATCG 2100
 TCAGCGGAA CATCAGGGAG AACATCTCA TGGGAGGCG ATATGACAA GCCCGATACC 2160
 TCAGGTGCT CCACTGCTGC TCCCTGAATC GGGACCTGGA ACTTCTGCC TTTGGAGACA 2220
 TGACAGAGT TGGAGAGCGG GGCCTCAACC TCTCTGGGG GCAGAAACAG AGGATCAGCC 2280
 TGGCCGCGC GACTTATCCG GACCGTCAGA TCTACCTGCT GGACGACCC CTGTCTGCTG 2340
 TGGACGCCCA CGTGGGGAG CACATTTTTC AGGAGTGCAT TAAGAAAGCA CTCAGGGGGA 2400
 AGACGCTGT CCGGTGAGC CACAGCTGCA AGTACTTAGA ATTTGTGGC CAGATCATTT 2460
 TGTGGGAAA TGGGAAAATC TGTGAAAATG GAACCTACAG TGAATTAATG CAGAAAAGG 2520
 GGAATATGC CCAACTTATC CAGAAGATGC ACAAGGAGC CACTTCGGAC ATGTTGCGG 2580
 ACACAGCAAA GATAGCAGAG AAGCCAAAGG TAGAAAGTCA GGCTCTGGCC ACCTCCCTGG 2640
 AAGAGTCTCT CAACGAAAT GCTGTGCCGG AGCATCAGCT CACACAGGAG GAGGAGATGG 2700
 AAGAAGGCTC CTGTGGTGG AGGGTCTACC ACCACTACAT CAGGCGAGCT GGAGGTTACA 2760
 TGTCTCTCTG CATATTTTTC TTCTCTGTGG TGCTGATGCT CTCTCTAACG ATCTTCAGCT 2820
 TCTGGTGGCT GAGCTACTGG TGGAGGAGG GCTCGGGGAC CAATAGCAGC CGAGAGAGCA 2880
 ATGGAACCAT GGCAGACCTG GGCACATTC CAGACAACTC TCAACTGTCC TTCTACCAGC 2940
 TGGTGTACGG GCTCAAGGCC CTGCTCTTCA TCTGTGTGGG GGTCTGCTCC TCAGGGATTT 3000
 TCACAAAGT CTCTGAGGAG GCATCCACGG CCGTGCACAA CAGGCTCTTC AACAGGTTT 3060
 TCCGCTGCC CATGAGTTTC TTGACACCA TCCCAATAGG CCGGCTTTTG AACTGCTTCG 3120
 CAGGGGACTT GGAACAGCTG GACCAGCTCT TGCCCATCTT TTCAGAGCAG TTCTGGTCC 3180
 TGTCTTAACT GGTGATGCGC GTCCGTGTGA TTGTGAGTGT GCTGTCTCCA TATATCTGT 3240
 TAATGGGAGC CATATTCATG GTTATTGCT TCAATTATTA TATGATGTTT AAGAAGGCCA 3300
 TCGGTGTGTT CAAGAGACTG GAGAAGTATA GCGGCTCTCC TTATTCTCC CACATCCTCA 3360
 ATTCTCTGCA AGGCTTGAGC TCCATCCATG TCTATGAAA AACTGAAGAC TTCTACAGCC 3420
 AGTTTAAGAG GCTGACTGAT GCGCAGATA ACTACCTGCT GTGTGTTCTA TCTTCCACAC 3480
 GATGGATGGC ATTGAGGCTG GAGATCATGA CCAACCTTGT GACCTTGGCT GTTGCCCTGT 3540
 TCGTGGCTTT TGGCATTTCC TCCACCCCT ACTCCTTAA AGTCATGGCT GTCAACATCG 3600
 TGCTGAGCT GCGCTCCAGC TTCCAGGCCA CTGCCCAGAT TGGCTTGGAG ACAGAGGCAC 3660
 AGTTCACGGC TGTAGAGAGG ATACTGAGT ACATGAAGAT GTGTGTCTCG GAAGCTCTCT 3720
 TACACATGGA AGGCACAAGT TGTCCCAGG GGTGGCCACA GCATGGGAA ATCATATTTT 3780
 AGGATTATCA CATGAATAC AGAGCAACA CACCCACGCT GCTTCAAGGC ATCAACCTGA 3840
 CCATCCGCGG CCACGAAGT GTGGGCTATG TGGGAAGGAC GGGCTCTGGG AAGTCTCTCT 3900
 TGGGCTGCTT CACTTCTCCG CTGGTGGAGC CCATGGCAG CCGGATCTCT ATTGAGCGCG 3960
 TGGACATTTG CAGCATGGCG CTGGAGGACT TGGGTTCGA GCTCTCAGTG ATCCCTCAAG 4020
 ATCCAGTGCT GCTCTCAGGA ACCATCAGAT TCAACCTAGA TCCCTTTGAC CGTCACACTG 4080
 ACCAGCAGAT CTGGGATGCC TTGGAGAGGA CATTOCTGAC CAAGGCCATC TCAAGTTTCC 4140
 CCAAAAGACT GCATACAGAT GTGTGTGAAA ACGGTGGGAA CTCTCTGTG GGGGAGAGGC 4200
 AGCTGCTCTG CACTTCTCCG GCTGTGCTTC GCAACTCCAA GATCATCCTT ATCGATGAAG 4260
 CCAACAGCTC CATGACATG GAGACAGACA CCGTATCCA GCGCACAATC CGTGAAGCCT 4320
 TCCAGGGCTG CACGCTGCTC GTCATTGCCC ACCGTGTAC CACTGTGCTG AACTGTGACC 4380
 ACATCCTGTT TATGGGCAAT GGAAGGTTGG TAGAATTTGA TCGGCCGGAG GTACTGCGGA 4440
 AGAAGCCTGG GTCAATTGTC GCGGCCCTCA TGGCCACAGC CACTTCTTCA CTGAGATAAG 4500
 GAGATGTGGA GACTTCATG AGGCTGGCAG CTGAGCTCAG AGGTTACAC AGGTGCAGCT 4560
 TCGAGGCCCA CAGTCTGCGA CCTTCTGTGT TGGAGATGAG AACTTCTCCT GGAAGCAGGG 4620
 GTAAATGTAG GGGGGGTGGG GATTGCTGGA TGGAAACCTT GGAATAGGCT ACTGTATGGC 4680
 TCTCAGACC TTAGAACCCC AGAACCATCT AAGACATGGG ATTCAGTATG CATGTGGTTC 4740
 TCTTTTAACT TTACATGCTG AATAATTTTA TAATAAGGTA AAGCTTATA GTTTTCTGAT 4800
 CTGTGTAGA AGTGTGCAA ATGCTGTACT GACTTGTAA AATAAAAC TAAGGAAAC 4860
 TC

Seq ID NO: 104 Protein sequence
 Protein Accession #: NP_149163.2

1 11 21 31 41 51
 65 MTRKRTYVWP NSSGGLVNRG IDIGDDMVSG LIYKTYTLQD GFWSQOERNP EAPGRAAVFP 60
 WGYDAALRT MIPFRPKPRF PAPQPLDNAG LPSYLVSWL TPLMIQSLRS RLDENTIPPL 120
 SVHDASDKNV QRLHRLWEEB VSRGIEKAS VLLVMLRFRQ TRLIFDALLG ICFCIASVLG 180
 PLLIIPKILE YSEEQLGNVV HGVGLCFALF LSECVKSLSF SSSNIIMORT AIRFRAAVSS 240
 FAFKLIQPK SVIHITSGEA ISPTGQDVNY LFEQVCYQPL VLITCASLVI CSISSYFIIG 300
 70 YTAFTIALCY LLVFLAVFM TRMAVKAQHR TSEVSDQRLR VTSEVLTCLK LIRMYIWEKP 360
 FAKIIEDLRR KERKLEKCG LVQSLTSITL FIIPTVATAV WVLIETSLKL KLTASMAFSM 420
 IASLNLRLS VFFVPIAVKG LTNSKSAVMR FRKFFLQESP VYVQTLQDP SKALVFEEAT 480
 LSNQQTCPGI VNGALELERK GBAEGMTRP RDALGPREEG NSLGPFLHKI NLVVSRCMML 540
 75 GVCNVTGSGK SSSLIAILEB MHLLEGSVGV QGSLAVFPQ ARIVSGNIRE NILMGAYDK 600
 ARYLQVLHCC SNRDLLELP FGDNTEIGER GLNLGGQXQ RISLARAVYS DRQIYLLDDP 660
 LSAVDARVGR HIFESCIKKT LRGKTVVIVT EQLQYLEFCG QIILLENGKI CENGTHSELM 720
 QKQKYAQLI QMHKEATS D MLQDTAKIAE KPKVESQALA TSLERSLNGN AVPEHQLTQE 780
 ERMEGSLRW RYVHYEQAA GGYMVSCILF FFFVLIVFLT IFSFNMWSYM LEQSGSINSS 840
 80 RESNGTMADL GNIDNPQLS FYQLVYGLMA LLLICVGVCS SGIFTKVIK ASTALHNKLF 900
 NKVRCFMSF FDTIPIGRLL NCFAGDLBQL DQLLPFBSQ FLVLSLMVIA VLLIVSVLSP 960
 YILLMGAIM VICFYIYMF KKAIGVFKRL ENYSRSLFS HILNLSQLGS SIRVYKTED 1020
 FISQFKRLTD AQNNYLLLEF SSTRWMLRL EIMINLVTLA VALFVAFGIS STPYSKVMA 1080
 VNIVLQLASS FQATARIGLE TEAQTAVER ILQYMKCVS EAPLHMEGTS CPQWPQHG 1140
 LIQDYHMKY RDNTPIVLHG INLTIRGHEV VGIVGRTGSG KSLGLMALFR LVEPMAGRIL 1200

IDGVDIQSIG LEDLRSLKV IPQDPVLLSG TIRFNLDPFD RHTDQOIWDA LERTFLTAKI 1260
 SKFPKLLHID VVENGGNFSV GERQLLCIAR AVLRNSKILL IDEATASIDM ETDLLIORTI 1320
 REAPQGCITVL VIAHRVITVL NCDHILVMGN GKVVEFDRPE VLRRKPGSLF AALMATATSS 1380
 LR

Seq ID NO: 105 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..4043

10 1 11 21 31 41 51
 ATGACTAGGA AGAGGACATA CTGGGTGCCC AACTCTTCCT GTGGCTCTGT GAATCGTGGC 60
 ATCGACATAG GCGATGACAT GGTTCAGGA CTATTATTATA AAACCTATAC TCTCCAGAT 120
 GGCCCTCTGA GTCAGCAAGA GAGAAATCTT GAGGCTCCAG GGAGGGCAGC TGTCCACCG 180
 TGGGGGAAGT ATGATGCTGC CTGAGAAACC ATGATTCCTT TCCGTCCCAA GCGGAGGTT 240
 CCTGCCCTCC AGCCCTCTGA CAATGCTGGC CTGTTCTCCT ACCTCACCGT GTCATGGCTC 300
 ACCCCCTCA TGATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCTCCACTG 360
 TCAGTCCATG ATGCTCTAGA CAAAAATGTC CAAAGGCTTC ACCGCTTTT GGAAGAAGAA 420
 GTCTCAAGGC GAGGGATGGA AAAAGCTTCA GTGCTTCTGT TGATGCTGAG GTTCCAGAGA 480
 ACAGAGTTGA TTTTGGATGC ACTTCTGGGC ATCTGCTTCT GCATTGOCAG TGTACTCGGG 540
 CCAATATTGA TTATACCAA GATCCTGGAA TATTCAGAAG AGCAGTTGGG GAATGTTGTC 600
 CATTGAGTGG GACTCTGCTT TGCCCTTTT CTCTCCGAAT GTGGAAGTTC TCTGAGTTTC 660
 TCCTCCAGTT GGATCATCAA CCAACGACCA GCCATCAGGT TCCAGCAGC TGTTCCTCTC 720
 TTTGCCCTTG AGAAGCTCAT CCAATTTAAG TCTGTAAATAC ACATCACCTC AGGAGAGGCC 780
 ATCAGCTTGA TTTTGGATGC TGTAAACTAC CTGTTTGAAG GGGGTGTGCTA TGGACCCCTA 840
 GTACTGATCA CTCTGCGATC GCTGGTCTAT TGCAGCAATT CTCTCACTT CATTTATGGA 900
 TACATGTCAT TTATTGTCAT CTTATGCTAT CTCTGCTT TCCACCTGCG GGTATTCATG 960
 ACAGAATGCG CTGTGAAGGC TCAGCATCAC ACATCTGAGG TCAGCGACCA GCGCATCCGT 1020
 GTGACCAGTG AAGTTCCTAC TTGCATTAAAG CTGATTAAAA TGTACACATG GGAGAAACCA 1080
 TTTGCAAAAA TCATTGAAGG TATGGAAAGT CTGACTTTCT GCTCCAAACG TGGTGTGGC 1140
 ATGGCTTCA GCAATGCTGC CTCCCTGAAT CTCTCTGGC TGTCACTGTT CTTTGTGCTC 1200
 ATGTGATCA AAGGCTCTAC GAATTCGAAG TCTGAGTGA TGAGGTTCAA GAAGTTTTC 1260
 CTCAGGAGA GCCCTGTTTT CTATGTCAG ACATTACAG ACCCAGCAA AGCTCTGCTC 1320
 TTTGAGGAGG CCACCTTGTC ATGGCAACAG ACCGTGTCGG GATCTGTCAA TGGGGCAGC 1380
 GAGCTGTGAG GAAACGGGCA TGCTTCTGAG GGGATGACCA GGCCTAGAGA TGCCCTCGGG 1440
 CCAGAGGAAG AAGGGAAACG CCTGGGCCCA GAGTTGCACA AGATCAACCT GGTGGTGTCC 1500
 AAGGGATCA TGTTAGGGGT CTGCGGCAAC ACBGGAGTG GTAGAGCAG CCTGTTGTCA 1560
 GCCATCTCTG AGGAGATGCA CTGTCTGAG GCTCTGCTGG GGTGACAGG AAGCCTGGCC 1620
 TATGTCCCCC AGCAGGCTTG GATCTGTGAG GGAACATCA GGGAGAACAT CCTCATGGGA 1680
 GGGCATATG ACAAGCCCG ATACCTCCAG GTGCTCCTCT GCTGCTCCTT GAATCGGGAC 1740
 CTGGAACCTC TGCCCTTTGG AGACATGACA GAGATTGGAG AGCGGGGCT CAACCTCTCT 1800
 GGGGGGCTCA AAGGCTCTAC CAGCTTGCC CGCGCGCTCT ATTCGACCTC TCAGATCTAC 1860
 CTGCTGGAGC ACCCCCTGTC TGCTGTGAGC GCCACGCTGG GGAAGCACAT TTTGAGGAG 1920
 TGCAATTAAGA AGACATCTAG GGGGAGAGCG GTCTCTCTGG TGACCCACCA GCTGCTGATC 1980
 TTAGAAATTT GTGGCCAGAT CATTTTGTGG GAAATGCGGA AATCTGTGA AATGGAATC 2040
 CACAGTGAAT TAATGAGAAA AAGGGGAAA TATGCCAAC TTATCCAGAA GATGCACAA 2100
 GAAGCCATCT GGTAGCTGTT GCAGGACACA GCAAGATAG CAGAGAAGCC AAGGTAGAA 2160
 AGTCAGGCTC TGCCACCTTC CTTGGAAGAG TCTCTCAAG GAAATGCTGT GCCGAGCAT 2220
 CAGCTCACAG AGGAGGAGGA GATGGAAGAA GGCTCTCTGA GTTGGAGGGT CTACCAACCA 2280
 TACATCCAGG CAGCTGGAGG TTACATGGTC TCTGTCAATA TTTCTCTCTT CGTGGTGTCT 2340
 ATGCTCTCTT TAACGATCTT CAGCTCTGAG TGGCTGAGCT ACTGGTTGGA GCAGGCTCTC 2400
 GGGACCAATA GCGAGGAGG GAGCAATGGA ACCATGGCAG AACTGGGCTA CATTCAGAC 2460
 AATCTCAAC TGCTCTCTTA CCAGCTGGTG TACGGCTCA ACCCCCTGCT CTCTATCTGT 2520
 GTGGGGGTCT GCTCTCTCAG GATTTTCACC AAAGTCAGGA GGAAGGCATC CACGGCCCTG 2580
 CACAAACAGC TCTTCAACAA GGTTTTCCGC TGCCCATGGA GTTCTCTTGA CACCATCCCA 2640
 ATAGGCCGGC TTTTGAAGCT CTTCGAGGG GACTTGAAC AGCTGGAACA GCTCTTGGCC 2700
 ATCTTTTCAG AGCAGTTCTT GGTCTGTGTC TTAATGGTGA TGCCCTGCTT GTTGATGTCT 2760
 AGTGTGCTGT CTCCATATAT CTTGTTAATG GGAGCCATA TCATGTTTAT TTGCTTCATT 2820
 TATTATATGA TGTTCAGAAA GGCCATCGGT GTGTTCAAGA GACTGGAGAA CTATAGCCCG 2880
 TCTCTTTTAT TCTCCACAT CTTCAATCTT CTGCAAGGCC TGAGCTCCAT CCATGTCTAT 2940
 GGAAGAACTG AAGACTCTAT CAGCCAGTTT AAGAGGCTGA CTGATGCGCA GAATAACTAC 3000
 CTGCTGTTGT TCTATCTTC CACAAGATGG ATGGCAATGA GGCTGGAGAT CATGAACCA 3060
 CTTGTGACCT TGGCTGTGTC CTTGTTGCTG GCTTTTGCCA TTTCTCTCAC CCCCCTACTC 3120
 TTTAAAGTCA TGGCTGTCAA CATGCTGCTG CAGCTGGCGT CAGCTTCCA GGCCTACTGC 3180
 CGATTGGCT TGGAGACAGA GGCACAGTTC ACGGCTGTAG AGAGGATACT GCAGTACATG 3240
 AAGATGTGTG TCTCGAAGC TCCTTTACAC ATGGAAGGCA CAAGTTGTCC CCAGGGGTGG 3300
 CCACAGCATG GGAATATCAT ATTTCAGGAT TATCAGATGA AATACAGAGA CAACACACCC 3360
 ACCGTGCTTC ACGGCATCAA CTTGACCATC CCGGCCACG AAGTGGTGGG CATGCTGGGA 3420
 AGGACGGGCT CTGGGAAGTC CTCTTGGGC ATGGCTCTCT TCCGCTCTGT GGAGCCCATG 3480
 GCAGGCCGGA TTCTCATTGA CGGCTGGAC ATTTGCAGCA TCGGCTTGGG GGACTTGGGG 3540
 TCCAGCTCTT CAGTATCTCC TCAAGATCCA GTGCTGCTCT CAGGAACCAT CAGATTCAAC 3600
 CTAGATCTCT TTGACCGTCA CACTGACAG CAGATCTGGG ATGCTTGGGA GAGGACATTC 3660
 CTGACCAAGG CCATCTCAAA GTTCCCAAAA AAGCTGCATA CAGATGTGCT GGAAGACGGT 3720
 GGAACCTTCT CTGTGGGGGA GAGGAGCTG CTCTGCATTG CCAGGGCTGT GCTTGCACAC 3780
 TCCAAGATCA TCCATTATCA TGAAAGCCAA GCTCCATG ACATGGAGAC AGACACCTCT 3840
 ATCCAGCGCA CAATTCGTGA AGCCTTCCAG GGCTGCACCG TGCTGCTCAT TGOCCACCGT 3900
 GTCAACACTG TGCTCAACTG TGACCATATC CTGGTTATGG GCAATGGGAA GGTGGTAGAA 3960
 TTTGATCGGC CGAGGTACT GCGGAAGAAG CTTGGGTCTT TGTTCGAGC CTCTATGGCC 4020
 ACAGCCACTT CTCTACTGAG ATAA

Seq ID NO: 106 Protein sequence
 Protein Accession #: Eos sequence

	1	11	21	31	41	51	
	MTKRKYVWVP	NSSGGLVNRG	IDIGDDMVSG	LIYKTYTLQD	GFWSQQERNP	EAPGRAAVFP	60
5	NGKYDAALRT	MIPFRPKPRP	PAPQFLDNAG	LPSYLTVSWL	TEPMIQSLRS	RIDENTIPPL	120
	SVDHSDKXNV	QRLHRLWESE	VRRGIERKAS	VLLVMLRFQR	THLIFDALLG	ICPCIASVLG	180
	PILIIPIKILE	YSEELQGNVV	HGVGLCFALP	LSECVKSLSP	SSSWIINQRT	AIRFRAAVSS	240
	FAFEKLIQFK	SVIHITSGEA	ISFPTGDVNY	LFEGVCYGPL	VLITCASLVI	CSISSYFIIG	300
	YTAFIALLCY	LLVFLPLAVM	TRMAVKAQHH	TSEVSDQRIK	VTSEVLTCIK	LKMYTWEKP	360
10	FAKIEGMS	LTFCSKPGDG	MAFSMLASLN	LLRLSVFFVP	IADVGLTNSK	BAVMRFKKFF	420
	LQESPVFYVQ	TLQDPSKALV	FERRATLSNQ	TCFPIVNGAL	ELERNHASE	GMTRPRDALG	480
	PSEBGNLSGP	ELKKNLVVVS	KGMMLGVCGN	TGSGKSSLLS	AILEEMELLE	GSVGVQGSIA	540
	YVPOQAWIVS	GNIRENILMG	GAYDKARYLQ	VLHCCSLNRD	LELLPPGDMT	EIGERGLNLS	600
	GGQKQRIsla	RAVYSDRQIV	LLDDPLSAVD	AHVGKHIFEE	CIKKTLRGKT	VVLVTHQLQY	660
15	LEPCQIILL	ENGKICENG	HSELWQKKGK	YAQLIQRMHK	EAISDMLQDT	AKIAEKPKVE	720
	SQALATSLER	SLNGNAVPEH	QLTQEEEMEE	BSLSWRVYHH	YIQAAGGYMV	SCIIFFFPVVL	780
	IVFLTIFSFV	WLSYWLQESG	GTNSRESNNG	TMADLGNIA	NPLSFPYQLV	YGLNALLLIC	840
	VGVCSGSIPT	KVTRKASTAL	HNKLFNKVFR	CPMSFFDTIP	IGRLNCFAG	DLEQLDQLLP	900
	IFSEQFVLVS	LMVIAVLLIV	SVLSPIYLLM	GAIIMVICFI	YMMFKKAIG	VEKRLNYSR	960
20	SPFLSHILNS	LQGLSSIHVY	GKTEDFISQF	KRLTDAQNNY	LLPLSSTRW	MALRLIIMTN	1020
	LVTIAVALFV	AFGISSTFYS	PKVMAVNIVL	QLASSFQATA	RIGLETAQF	TAVERILQYM	1080
	KMCVSEAPLE	MEGTSCFPQW	PQHGLIIFQD	YHMKYRDNTP	TVLHGINTLT	HGHEVVGIVG	1140
	RTGSGKSSLG	MAIFRLVERM	AGRILIDGVD	ICSIGLEDLR	SKLSVIPQDE	VLLSGTIRFN	1200
	LDPPDRHTDQ	QIWDALERTF	LTKAISKFPK	KLHTDVVENG	GNFSVGERQL	LCIARAVLRN	1260
25	SKILLIDEAT	ASIDMSTDTL	IQRTIREAPQ	GCTVLVIAHR	VITVLNCDHI	LVMGNGKVVE	1320
	FDRPEVLRRK	PGSLPAALMA	TATSSLR				

Seq ID NO: 107 DNA sequence
Nucleic Acid Accession #: NM_024022
Coding sequence: 202..1563

	1	11	21	31	41	51	
	ACCGGCGACC	GGACGGCTCG	GGTACTTTCG	TTCTTAATTA	GGTCATGCCC	GTGTGAGCCA	60
35	GGAAAGGGCT	GTGTTTATGG	GAAGCCAGTA	ACACTGTGGC	CTACTATCTC	TTCGTGGTGT	120
	CCATCTACAT	TTTTGGGACT	CGGGAATTAT	GAGGTAGAGG	TGGAGGCGGA	CGCGGATGTC	180
	AGAGGTCTCT	AAATAGTCAC	CATGGGGGAA	AATGATCCGC	CTGCTGTGTA	AGCCCCCTTC	240
	TCATTCCGAT	CGCTTTTGGG	CCTTGATGAT	TTGAAAATAA	GTCTGTGTC	ACCCAGATGCA	300
	GATGCTGTTG	CTGCACAGAT	CCTGTCACTG	CTGCCATGTA	AGTTTTTTC	AATCATCGTC	360
40	ATTGGGATCA	TGTCATTGAT	ATTAGCACTG	GCCATTGCTC	TGGGCATCCA	CTTCCACTGC	420
	TCAGGGAAGT	ACAGATGTCG	CTCATCCTTT	AAGTGTATCG	AGCTGATAGC	TCGATGTGAC	480
	GGAGTCTCGG	ATTGCAAGA	CGGGGAGGAC	GAGTACCCCT	GTGTCCGGGT	GGTGTGTGAG	540
	AATGCCGTGC	TCCAGGTGTT	CACAGCTGCT	TCGTGGAAGA	CCATGTGCTC	CGATGACTGG	600
	AAGGGTCACT	ACGCAAAATG	TGCCGTGTC	CAACTGGGTT	TCCCAAGCTA	TGTGAGTTCA	660
45	GATAACCTCA	GAGTGAGCTC	GCTGGAGGGG	CACTTCCGGG	AGGACTTTGT	GTCCATCGAT	720
	CACCTCTTGC	CAGATACCAA	GGTGACTGCA	TTACACCACT	CAGTATATGT	GAGGGAGGGA	780
	TGTGCTCTTG	GCCAGTGGT	TACCTTGCG	TGCACAGCCT	GIGGTCTAG	AAGGGGCTAC	840
	AGCTCAOCCA	TCGTGGGTGG	AAACATGTCT	TTGCTCTGCG	AGTGGCCCTG	GCAGGCCAGC	900
	CTTCAGTTCC	AGGGCTACCA	CCTGTGCGGG	GGCTCTGTCA	TCAGCCCTCT	GTGGATCATC	960
50	ACTGCTGCAC	ACTGTGTTTA	TGACTTGTAC	CTCCCCAAGT	CATGACCAT	CCAGGTGGGT	1020
	CTAGTTTCCC	TGTTGGACAA	TCCAGCCCCA	TCCCACCTGG	TGGAGAAGAT	TGTCTACCAC	1080
	AGCAAGTACA	AGCCAAAGAG	GCTGGGCAAT	GACATCGCCC	TTATGAAGCT	GGCCGGGCCA	1140
	CTCACGTTCA	ATGAAATGAT	CCAGCCCTGT	TGCCCTGCCA	ACTCTGAAGA	GAACCTCCCC	1200
	GATGGAAGAG	TGTGCTGGAC	GTCAGGATGG	GGGGCCACAG	AGGATGGAGG	TGACGCTCC	1260
55	CCGTGCTGTA	ACCAGCGGCG	GGTCCCTTTG	ATTTCACACA	AGATCTGCAA	CCACAGGGAC	1320
	GTGTADGGTG	GCATCATCTC	CCCTCCATG	CTCTGCGCGG	GCTACCTGAC	GGGTGGCGTG	1380
	GACAGCTGCC	AGGGGGACAG	CGGGGGGCCC	CTGGTGTGTC	AAGAGAGGAG	GCTGTGGAGG	1440
	TTAGTGGGAG	CGACAGCTTT	TGGCATCGGC	TGGCGAGAGG	TAACAAGGCC	TGGGTGTGAT	1500
	ACCTGTGTCA	CCTCCTTCCT	GGACTGGATC	CACGAGCAGA	TGGAGAGAGA	CCTAAAAACC	1560
60	TGAAGAGGAA	GGGACCAAGT	AGCCACCTGA	GTTCCTGAGG	TGATGAAGAC	AGCCGATCTC	1620
	TCCCTCTGAC	TCCCGTGTAG	GAACCTGCAC	ACGAGCAGAC	ACCCTTGGAG	CTCTGAGTTC	1680
	CGGCACCAAT	AGCAGGCCCC	AAAGAGGCAC	CCTTCCATCT	GATTCCAGCA	CAACCTTCAA	1740
	GCTGCTTTTT	GTTTTGTGTT	TTTTTGAGGT	GGAGTCTGCG	TCGTGTGCCC	AGGCTGGAGT	1800
	GCAGTGGCGA	AATCCCTGCT	CACCTGCAGC	TCCCTTCCCT	TGGTTCAGAG	GATTCTCTTG	1860
65	CCTCAGCTTC	CCCATGAGCT	GGGACCACAG	GTGCCCGCCA	CCACACCCAA	CTAATTTTGT	1920
	TATTTTGTAG	AGAGACAGGG	TTTCACCATG	TTGGCCAGGC	TGCTCTCAA	CCCCTGACCT	1980
	CAATGATGT	GCCTGCTTCA	GCTCCTCCCA	GTGCTGGGAT	TACAGGCATG	GGCCACCCAG	2040
	CCTAGCCTCA	CGCTCCTTTC	TGATCTTCAC	TAAGAACAAA	AGAAGCAGCA	ACTTGCAAGG	2100
	GGCGCCCTTC	CCACTGGTCC	ATCTGGTTT	CTCTCCAGGG	GTCTTGCAA	ATTCCTGACG	2160
70	AGATAAGCAG	TTATGTGACC	TCACGTGCAA	AGCCACCCAC	AGCCACTCAG	AAAAGACGCA	2220
	CCAGCCCGAG	AGTGCAGAAC	TGCAGTCACT	GCACGTTTTC	ATCTCTAGGG	ACCAGAACCA	2280
	AACCCACCTG	TTTACTTTC	AAGACTTATT	TTACATGTG	GGGAGGTTAA	TCTAGGAATG	2340
	ACTGTTTAA	GGCCTTATT	CATGATTCT	TTGTAGCATT	TGTTGCTTGA	CGTATTATG	2400
75	TCCTTTGATT	CCAAATAATA	TGTTTCTTTC	CCTCAAAAAA	AAAAA	AAAAA	2460

Seq ID NO: 108 Protein sequence
Protein Accession #: NP_076927

	1	11	21	31	41	51	
	MGENOPPAVE	APPSFRLSFG	LDLKLISFVA	PDRDAVAQI	LSLLPLKFPF	IIVIGIILI	60
80	LALALGLGIH	PDCSGKYRCR	SSFKCIELIA	XCDSVSDCKD	GEDEYRCVRV	GGQNAVQLQV	120
	TAASWKTMC	DDWNGHYANV	ACQLGFPSY	VSSDNLRVSS	LEGQFRREFV	SIDHLLPDDK	180
	VTALHHSVYV	REGCASGHVV	TLQCTACGER	RGYSSRIVGG	NMSLLSQWPN	QASLQFQGXH	240

LOGG8VITPL WIITAAHCYV DLYLPKSWTI QVGLVSLIDN PAPSHLVEKI VYHSKYKPKR 300
 LGNDIALMKL AGPTTFNEMI QPVCLPNSEK NFPDGVKVCWT SGWGATEDGG DASPVLNHAA 360
 VPLISNKKCN HRDVYGGIIS PSMCLAGYLT GGVDSCQGDG GGLVQCQERR LWKLVGATSF 420
 GIGCAEVNKP GUYTRVTSFL DMTHQMERD LKT

Seq ID NO: 109 DNA sequence
 Nucleic Acid Accession #: NM_000493.2
 Coding sequence: 97..2139

1 11 21 31 41 51
 | | | | |
 CACCTTCTGC ACTGCTCATC TGGGCAGAGG AAGCTTCAGA AAGTGCCAA GGCACCATCT 60
 CCAGGAACCTC CCAGCACGCA GAATCCATCT GAGAATATGC TGCCACAAAT ACCCTTTTGT 120
 CTGCTAGTAT CCTTGAACCTT GGTTCATGGA GTGTTTTACG CTGAACGATA CCAATGCCCC 180
 ACAGGCATAA AAGGCTTACC ACCCAACACC AAGACACAGT TCTTCATTCC CTACACCATC 240
 AAGAGTAAAG GTATAGCAGT AAGAGGAGAG CAGAGTACTC CTGGTCCACC AGGCCCTGCT 300
 GGACCTCGAG GGCACCCAGG TCCTTCTGGA CCACCAGGAA AACCAGGCTA CGGAAGTCCT 360
 GGACTCCAAG GAGAGCCAGG GTTGCCAGGA CTACCGGGAC CATCAGCTGT AGGGAACCA 420
 GTGTGTCCAG GACTCCAGG AARACCCAGG GAGAGAGGAC CATATGGACC AAAAGGAGAT 480
 GTTGAACCAAG GTGCTCAGG AGGACCCCGG GCGCCACCGG GAACACCTGG AATCCCTGGA 540
 CCGGCTGGAA TTCTGTGTCC AGGAAACCTT GGACACAGG GACCCACAGG AGCCCCAGGA 600
 CCGAGGGGCT TTCTGTGAGA AAAGGGTGCA CCAGGAGTCC CTGGTATGAA TGGACAGAAA 660
 GGGGAATAGG GATATGGTGC TCCTGGTGTG CCAGGTGAGA GGGGTCTTCC AGGCCCTCAG 720
 GGTCCACAGG GACCATCTGG CCCTCCTGGA GTGGGAAAAA GAGGTGAAAA TGGGGTTCCA 780
 GGACAGCCAG GCATCAAGG TGATAGAGGT TTTCCGGGAG AAATGGGACC AATGGGCCCA 840
 CCAGGTCCCC AAGGCCCTCC TGGGGAACCA GGGCCAGAGG GCATTGGAAG GCCAGGAGCT 900
 CGTGAAGCCC CCGCTCAGG AGGGATTCCA GGAACAAAG GTCTCCTTGG GGCTCCAGGA 960
 ATAGCTGGGC CCCCAGGGCC TCCTGGCTTT GGGAAACCA GCTTGCCAGG CCTGAAGGGA 1020
 GAAAGAGCAG CTGCTGGCCT TCCTGGGGGT CCAGGTGCCA AAGGGGAACA AGGGCCAGCA 1080
 GGTCTTCTCT GGAAGCAGG TCTGACTGGA CCCCCTGGGA ATATGGGACC CCAAGGACCA 1140
 AAAGGATACC CCGGTAGCCA TGCTCTCCCA GGCCTTAAAG GTGAGACAGG GCCAGCTGGG 1200
 CCTGCAGGAT ACCCTGGGGC TAAGGTTGAA AGGGGTCCCT CTGGGTGAGA TGGAAAAACA 1260
 GGGTACCCAG GAAAACCAAG TCTCGATGGT CCTAAGGGTA ACCCAGGGTT ACCAGGTCCA 1320
 AAAGGTGATC CTGAGGTGG AGGACCTCCT GGTCTCCAG GCCCTGTGGG CCCAGCAGGA 1380
 GCAAGGGGAA TGCCTCGACA CAATGGAGAG GCTGGCCCAA GAGGTGCCCC TGGGAATACA 1440
 GGTACTAGAG GCCTATATGG GCCACCAAGC ATTCAGGAT TCCTTGGGTC TAAAGGGGAT 1500
 CCAGGAAGTC CCGTCTCTCC TGGCCAGCT GGCATAGCAA CTAAGGGCTT CAATGGACCC 1560
 ACCGGGCCAC CAGGGCTTCC AGGTCCAGA GGCACCTCTG GAGAGCCTGG TCTTCCAGGG 1620
 CCCCCTGGGC CTCACGGCCC ACCAGGTCAA GCAGTCTATG CTGAGGGTTT TATAAAGGCA 1680
 GGCCAAAGGC CCAGCTCTTC TGGGACCCCT CTGTGTAGTG CCAACCAAGG GGTAAACAGA 1740
 ATGCTGTGTT CTGCTTTTAC TGTTATTCTC TCCAAAGCTT ACCCAGCAAT AGGAACTCCC 1800
 ATACCATTTG ATAAATTTT GTATAACAGG CAACAGCATT ATBACCCAAG GACTGGAATC 1860
 TTACTTGTGC AGATAACAGG AATATACTAT TTTTCATACC ACGTGCATGT GAAAGGGACT 1920
 CATGTTTGGG TAGGCTGTGA TAAGATGGGC ACCCTGTAA TGTACACCTA TGTATATAC 1980
 ACCAAAGGCT ACCTGGATTA GGCCTCAGGG AGTGCCATCA TCGATCTCAC AGAAATGAC 2040
 CAGGTGTGGC TCCAGCTTCC CAATGCCAGG TCAATGGCC TATACTCTTC TGASTATGTC 2100
 CACTCCTCTT TCTCAGGATT CCTAGTGGCT CCAATGTGAG TACACCCAC AGAGCTAATC 2160
 TAAATCTTGT GCTAGAAAAA GCATTCTCTA ACTCTACCCC ACCCTACAAA ATGCATATGG 2220
 AGGTAGGCTG AAAAGATATG AATTTTATTT TTCTGAAATA CAGATTGAG CTATCAGACC 2280
 AACAAACCTT CCCCCTGAAA AGTGAGCAGC AACGTAAAAA CGTATGTGAA GCCTCTCTTG 2340
 AATTTCTAGT TAGCAATCTT AAGGCTCTTT AAGGTTTCT CCAATATTAA AAAATATCAC 2400
 CAAAGAGTC CTGCTATGTT AAAAAACAAC AACAAAAAAC AAGCAACAA AAAAAAAT 2460
 TAAAAAATA AACAGAAATA GAGCTCTAAG TTATGTGAAA TTTGATTGGA GAAACTCGGC 2520
 ATTTCTTTT TAAAAAGGCC TGTTTCTAAC TATGAATATG AGAATCTTA GGAACATCC 2580
 AGGAGGTATC ATATAACTTT GTAGAACTTA AATACTTGAA TATTCAAAAT TAAAAGACAC 2640
 TGTATCCCTT AAAATATTTT TGATGGTGCA CTACTCTGAG GCGGTATGAG CCCTTCTCAT 2700
 CAATATCTAT TCAATATATC AGGTGCATAT ATACTTGTA AAGCTCTTAT ATAAAAAAGC 2760
 CCAAAATAT TGAAGTTAT CTGAAATGCA AGGTGCTTTC ATCAATGAAC CTTTCAAAA 2820
 CTTTCTATG ATGTCAGAGA AGCTTTTAT ATACCCAGCA TAAGTTGAAA ACAGGTATCT 2880
 GAOCATATCT TATTTAGTTA ACACAAGTGT GATTAATTTG ATTTCTTTAA TTCTTATTG 2940
 AATCTTATGT GATATGATT TCTGGATTTA CAGAACATTA GCACATGTAC CTGTGCTCTC 3000
 CCATTCAAGT GAAGTTATAA TTTACCTGA GGGTTTCAAA ATTGCACTAG AAGTGGAGAT 3060
 ATATTATTTA TTTATGCACT GTACTGTATT TTTATATTGC TGTTTAAAC TTTTAAGCTG 3120
 TGCTCACTT ATTAAGCAC AAAATGTTT ACCTACTCCT TATTTACGAC ACATTAAT 3180
 AACATCAATA GATTTTAAAG CTGAATTAAT TTGAAAGCAG CAATTGTCTG TTCTCAACCA 3240
 TTCTTCAAG GCTTTTCAAT CGACACAATA AAATAACATC AATAG

Seq ID NO: 110 Protein sequence
 Protein Accession #: NP_000484.2

1 11 21 31 41 51
 | | | | |
 MLPQIPFLLL VSLNLVHGVF YAEYQMPTG LKGFLENTKI QFFIPTYIKS KGLAVRGRQG 60
 TPQPPGPGAG RGHGPGSPGF GKPGYGSPL QGEPGLPGPF GPSAVGKPGV PGLPGKPGER 120
 GPYGPKGDVG EAGLPGPRGF PGPPGIPGPA GISVPGKPGQ QGPTGAPGPR GPPGKGPAGP 180
 VPGMNGQKGE MYGAGPRPG ERLGPGPQSF TGPSPGPFVGV KRSENGVPGQ PGIKGDRGFP 240
 GEMGPIGPGP PQGPPGSRGP EGIGKPGAAG APGQPGIIFGT KGLPGAPGIA GPPGPPGPGK 300
 PGLPGLKGER GPAGLPQPGP ARGEQGEAGL PGKPLGTGPP GMMGPQGPKG IPGSEGLPGP 360
 XGETGPAGPA GYPAKGERG SPGSDGKPGY PGKPLDGPK CNPGLPGPKG DPGVCGPRGL 420
 PGVPGAGAK GMPGHNGEAG PRGAPGIPGT RGPIGPPGIP GFFGSKGDPG SPGPPGPGAI 480
 ATKGLNGTPT PPGPPGPRGH SGEPGLPGPF GPPGPPGQAV MPEGFIAKAG RPSLGGTFLV 540
 SANQGVGTMP VSAFTVILSK AYPAGTPIP FDKILYNROQ HYDFRTGIFT CQIFGIYFFS 600
 YHVHVKGTHV WVGLYKNGTP VMYTYDEYTK GYLDQAGGSA IIDLTENDQV WLQLENAESN 660
 GLYSSEYVHS SPGFLVAPM

Seq ID NO: 111 DNA sequence
Nucleic Acid Accession #: NM_000949
Coding sequence: 285..2153

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

1 11 21 31 41 51
GGAGGCTGAA ATCCCCAGAC GCCGGTTTTC TGGGCTGGGC TTTCTGCTTA CTCACCTCTT 60
CTCCCTCTTT CTGGATTTTA COGACCGTTC GCGAAACAGC TTTCACACCA ATGGAGCTTC 120
ATGCTCTCGT GCAGGAAGTA CTCATCGACT GATGTGGCAG ACTTTGCTCC CTGACAAAAC 180
TAAAGAATCT TCTATTTCAT GGAGGCGAAG ACTGAGGATG CTTTCCACAT GAAACCTGAA 240
GTGAACCTCT GATACATTTC CTGCAGCAAG AGAAGGCAGC CAACATGAAG GAAATGTGG 300
CATCTGCAAC CGTTTTCAC TCGCTACTTT TTCTCAACAC CTGCCTTCTG AATGGACAGT 360
TACCTCTCTG AAAACCTGAG ATCTTTAAAT GTCGTTCTCC CAATAGGAA ACATTACCT 420
GCTGGTGGAG GCTCGGGACA GATGGAGGAC TTCTTACCAA TTATTCACTG ACTTACCACA 480
GGGAAGGAGA GACACTCATG CATGAATGTC CAGACTACAT AACCGGTGGC CCCAACTCCT 540
GCCACTTTGG CAGACAGTAC ACCTCCATGT GGAGGACATA CATCATGATG GTCATGCCA 600
CTAACAGATG GGGAGCAGT TTCTGGGATG AACCTTAATG GGACGTGACT TACATAGTTC 660
AGCCAGACCC TCCCTTGGAG CTGGCTGTGG AAGTAAACA GCCAGAAGAC AGAAACCCCT 720
ACCTGTGGAT TAAATGGTCT CCACCTACCC TGATTGACTT AAAAACTGGT TGGTTACGCG 780
TCTGTATAGA AATTCTGATTA AAACCCGAGA AAGCAGCTGA GTGGGAGATC CATTTGCTG 840
GGCAGCAAA AGATTTTAAG ATTCTCAGCC TACATCCAGG ACAGAAATAC CTGTCCAGG 900
TTCGCTGCAA ACCAGACCAT GATACTGGA GTGCATGGAG TCCAGCGACC TTCATTGAGA 960
TACCTAGTGA CTTACCATG AATGATACAA CCGTGTGGAT CTCGTGGGCT GTCCTTTCTG 1020
CTGTCTCTG GCTCGGGACA GTCTGGGAG TGGCTTTGAA GGGCTATAGC ATGGTGAACCT 1080
GCATCTTCCC GGCAGTCTCT GGGCCAAAAG TAAAGGATT TGTATGCTAT CTGTGGGAGA 1140
AGGGCAAGTC TGAAGAACTA CTGAGTGCTT TGGGATGCCA AGACTTTCTT CCCACTTCTG 1200
ACTATGAGGA TTGCTGGTG GAGTATTTAG AAGTAGATGA TAGTGAGGAC CAGCATCTAA 1260
TGTCACTCCA TTCAAAAGAA CACCCAGATC AAGGTATGAA ACCCAGATAC CTGGATCTCT 1320
ACACTGACTC AGCCCGGGGG AGCTGTGACA GCGCTTCCCT TTGTCTGAA AAGTGTGAGG 1380
AACCCAGGCG CATCTCTCC ACATTCTATG ATCTGAGGT CATTGAGAAG CCAGAGAATC 1440
CTGAACCAAC CCACACCTGG GACCCCCAGT GCATAAGCAT GGAAGGCAAA ATCCCTATT 1500
TTCATGCTGG TGGATCCAAA TGTTCACAT GGCCTTACC ACAGCCGAGC CAGCACAACC 1560
CCAGATCTCT TIACCACAAT ATTACTGATG TGTGTGAGCT GGCCTGTGGC CCGTGCAGGTG 1620
CACCGGAGCT TCTGTGAAT GAAGCAGGTA AAGATGCTTT AAAATCCTCT CAACACATTA 1680
AGTCTAGAGA AGAGGGAAAG GCAACCCAGC AGAGGGAGGT AGAAGGCTTC CATCTGAGA 1740
CTGACCAGGA TACGCCCTGG CTGCTGCCCC AGGAGAAAAC CCGCTTTGGC TCGCTAAAC 1800
CCTTGGATTA TGTGGGATT CACAAGGTCA ACAAGATGG TGCATTATCA TTGCTACCAA 1860
AACAGAGAGA GACAGCGCGC AAGCCCAAGA AGCCCGGAGC TCCTGAGAAC AATAAGGAGT 1920
ATGCGCAAGT GTCCGGGGTC ATGATAAACA ACATCTGGT GTTGGTGCCA GATCCACATG 1980
CTAAAAAGCT GGTCTGCTTT GAAGAATCAG CCAAGAGGCG CCCACCATCA CTGAAACAGA 2040
ATCAAGCTGA GAAAGCCCTG CCAACTTCA CTGCAACATC AAGCAAGTGC AGGCTCCAGC 2100
TGGGTGGTIT GGAATTACCTG GATCCCGCAT GTTTTACACA CTCCTTTTAC TGAAGCTTG 2160
ACTAATGGA TATTTGTTTA AATGTGATG TTCTTTCAGG TAACACTACA GAGTACGTGA 2220
AATGCTCAAG AATGTGATCA GACTGACACT ACTAAGCTC CCAGCTCCTT TCATGCTCCA 2280
TTTTTAACCA CTGCTCTCT TCTCCAGCAG CTGATTCAG AACAAATCAT TATGTTTCT 2340
AAGTGTGAT TGTAGATTA CTTTGTGCTG TTAGTTATAA AACTATGTGT TCAATGAAAT 2400
AAAAGCACAC TGCTTAGTAT TCTTGAGGGA CAATGCCAAT AGGTATATCC TCTGGAAAAG 2460
GCTTTTATGA TTGGCATGG GACAGACGGA AATGAAATG TCAGAAATGT TTACCATAGA 2520
AAGATGACAA AAGAAAATTT TCACATAGG AATATGCCAT GAAATTTGCT TTTGAAAAAC 2580
AACTGCATAA CTTTACACT CTTGCTCCAT TTTATTAGGA TTACCCAAAT ATAACCATTT 2640
AAGAAAGAA TGCATTCCAG AACAAATGT TTACATAAGT TCCTATACCT TACTGACACA 2700
TTGCTGATAT GCAAGTAGA AAT

Seq ID NO: 112 Protein sequence
Protein Accession #: NP_000940

60
65
70
75
80

1 11 21 31 41 51
MKENVASATV FTLLFLNLTC LLNGQLPFGK PEIFKCRSPN KEIFTCWWRP GTDGGLEPTNY 60
SLTYHREGET LMBBCPDYIT GGPNSCHFGK QYTSMTYTI MMVNATNQMG SSFSEDELYVD 120
VTYIVQEDPP LELAVEVKQP EDRKPYLNIK WSPPTLIDLK TGNFTLLYEI RLKPEKAAEW 180
EIHFGAQOTE FKILSLHPGQ KYLVQVRCKP DEGYNSAWSP ATFIQIFBDF TMNDITWVIS 240
VAVLSAVICL IIVWAVALKG YSMVTCIFPP VPGFKIKGFD AHLLEKSKSE KLLSALGQDD 300
FPPTSDYEDL LVEYLEVDDG EDQHLMSVES KEHPSQGMKP TYLDPTDSDG RGSCEPSLL 360
SEKCEPQAN PSTFYDPEVI EKFPNPETH TWDPQCLIME GKIFYFHAGG SKCSTWPLPQ 420
PSQHNPRSSY HNTIDVCELA VGEAGAPATL LNEAGKDALK S9QTIKSREE GKATQOREVE 480
SFSEYDQDT PHLLPQEKTP FGSAPLDYV EIRKVNKDA LSLLPKQREM SGPICKPGTP 540
ENNKEYAKVS GYMDNNILVL VPDPAKNVA CFRESAKEAP PSLBNQAEK ALANFTATSS 600
KCRQLGLGLD YLDPACFTHS FH

Seq ID NO: 113 DNA sequence
Nucleic Acid Accession #: XM_062811
Coding sequence: 1..888

75
80

1 11 21 31 41 51
ATGTGGGGGG CTGCGCGCTC GTCCGCTCTC TCATCTGGA ACGCCGCTTC GTCCTGCGAG 60
CTGCTGCTGG CTGCGCTGCT GCGCGCGGGG GCGAGGGCCA GCGGCGAGTA CTGCCACGGC 120
TGGCTGGAGC CGCAGGCGCT CTGGCGCATC GCGTTCAGT GTCCCGAGCG CTTGCAAGGC 180
GGCGAGGCCA CCACTGCTGC CGGCAGCTGC GCGTTCGCT ACTGCTGCTC CAGCGCGGAG 240
GCGCGCTGG ACACAGCGCG CTGCGACAA GACCGCCAGC AGGGCGCTGG CGAGCTGCGC 300
CGGGCGGACA AAGACGCGCC CGACGCTCG GCAATGCCCA TCTACGTGCT GTTCTCATTT 360
GTTGCTCTCG TGTGTGCTGC CTTTATCATC TTGGGGTCCC TGGTGGCAGC CTGTTGCTGC 420

AGATGTCTCC GGCCTAAGCA GGATCCCCAG CAGAGCCGAG CCCCAGGGGG TAACCGCTTG 480
 ATGGAGACCA TCCCCATGAT CCCCAGTGCC AGCACTCCCC GGGGGTCGTC CTCACGCCAG 540
 TCCAGCACAG CTGCCAGTTC CAGCTCCAGC GCCAAGCTCAG GGGCCCCGGC GCCCCCAACA 600
 AGGTACACAGA CCAACTGTTG CTTCGCCGAA GGGACCATGA ACACACTGTA TGTCAACATG 660
 CCCCAGCAAT TCTCTGTGCT GAAGTGTGAG CAGGCCACCC AGATTGTGCC ACATCAAGGG 720
 CAGTATCTGC ATCCCCATA CGTGGGGTAC ACCTGTCAGC ACGACTCTGT GCCCATGACA 780
 GCTGTGCCAC CTTCATGGA CGGCCTGCA GCTGGCTACA GGCAGATTCA GTCCCCCTTC 840
 CCTCACACCA ACAGTGAACA GAAGATGTAC CCAGCGGTGA CTGTATAA

Seq ID NO: 114 Protein sequence
Protein Accession #: XP_062811

1 11 21 31 41 51
 MNGARRSSVS SSWNAASLLQ LLLAALLAAG ARASGEYCHG WLDAQGVWRI GPQCPERFDG 60
 GDATICCGSC ALRYCCSSAE ARLDQGGCDN DRQQGAGEPG RADKDGPDS AVPIYVPFLI 120
 VGSVFFAFII LGSILVAACCC RCLRPKQDFQ QSRAPGGNRL METIPMIPSA STSRGSSSEQ 180
 SSTASSSSSS ANSGARAPPT RQTNCCLEPE GTMNNVYVNM PTNFSVLNCQ QATQIVPHQG 240
 QYLHFFPVGY TVQHSVPMT AVPPFMDGLQ PGYRQIQSPF PHTNSEQKMY RAVTV

Seq ID NO: 115 DNA sequence
Nucleic Acid Accession #: NM_013257
Coding sequence: 223..1512

1 11 21 31 41 51
 GGTGTGCTCT TGAGGGGATTA AATGCAAAAG GATCACACCA TGGACTACAA GGAAAGCTGC 60
 CCAAGTGTAA GCATTCCAG CTCCGATGAA CACAGAGAGA AAAAGAGAG GTTTACTGTT 120
 TATARAAGTTC TGGTTTCAGT GGAAGAAGT GAATGGTITG TCTTCAGGAG ATATGCAGAG 180
 TTTGATAAAC TTTATAACAC TTTAAAAAAA CAGTTTCTTG CTATGGCCCT GAAGATTCCCT 240
 GCCAAGAGAA TATTTGGTGA TATTTTGTAT CCAGATTTTA TTAACAAAG ACGAGCAGGA 300
 CTAAACGAAT TCATTAGAA CCTAGTTAGG TATCCAGAAC TTTATAACCA TCCAGATGTC 360
 AGAGCATTC TCAAAATGGA CAGTCCAAA CACCACTCAG ATCATCTGA AGATGAGGAT 420
 GAAAGAAGTT CTCAGAAGCT ACACCTCACC TCACAGAACA TCAACCTGGG ACCGTCTGGA 480
 AATCCTCATG CCAAAACCAAC TGACTTTGAT TTCCTAAAAA TTATTGGAAA AGGCAGCTTT 540
 GGCAAGGTTT TTTCTGCAAA ACGGAAACTG GATGGAAAAT TTTATGCTGT CAAAGTGTTA 600
 CAGAAAAAAA TAGTTCTCAA CAGAAAAGAG CAAAAACATA TTATGGCTGA ACGTAATGTG 660
 CTCCTGAAA ATGTGAAACA TCCGTTTTTG GTTGGATTGC ATTATTCCTT CCAACAACCT 720
 GAAAGCTTT ATTGTGTTCT GGAATTTGTT AATGGAGGG AGCTTTTTTT CCACCTACAA 780
 AGAGAACGGT CCTTTCCTGA GCACAGAGCT AGGTTTTACG CTGCTGAAT TGTCTAGTGA 840
 TTGGGTACT TACATTCAT CAAATAAGTA TACAGAGACT TGAACACAGA AATATTCCT 900
 TTGGATTCAG TAGACATGT TGTCTTAACA GATTTTGGGC TTTGTAAGA AGGAATGTCT 960
 ATTTCTGACA CCACTACCA ATTTTGTGG ACACCAAGAT ATCTTGCACT TGAAGTAAT 1020
 AGAAACAGC CATTAGCAAA TCTGTAGAT TGGTGGTGCC TTGGGCTGT TCTGTATGAA 1080
 ATGCTGTATG GATTGCTTCC TTTTATATGC CGAGATGTTG CTGAATGTA TGACAATATC 1140
 CTTCAACAA CCAATAGTTT GAGGCCAGGA GTGAGTCTTA CAGCTGGTTC CATCTGGAA 1200
 GAATCTCTAG AAAAAGACAG GCAAAATGGA CTGGTGCCA AGGAAGACTT TCTTGAAT 1260
 CAGATCTATC CTTTCTTGA ATCACTCAGC TGGGCTGACC TTGTACAAA GAAGATTCCA 1320
 CCACCATTTA ATCTAATGT GGCCTGACCA GATGATATCA GAACTTTGA CACGACATT 1380
 ACAGAGAAA CAGTCCATA TTTCTGTGT GTATCTCTG ACTATTCTAT AGTGAATGCC 1440
 AGTGTATTG AGGCAGATGA TGCATTGCT GGTTCCTCT ATGCACCTCC TTCAGAGAC 1500
 TATTTTTTGT GAGCATTTG CCATTAGAA ACCATTGAGC AAAATAAGTC TATAGATGG 1560
 ACTGAACTT CATTTGTGT GAATATATC AATATGTAT AACTAGTGCC TCATTTTAT 1620
 ATGTAATGAT GAAATATG AAAAAATGA TTTCTCTTA TGTGCAAGAA AATAGGGCA 1680
 TTTCAAGAG CAGTTTTGAT TAAATTTAT ATTCTGTCTT AATAAGCTTA TTTTAAACA 1740
 ATTTAAAGC TATTATCTT AGCATTAACC TATTTTAAA GAAACCTTT TTGCTATTGA 1800
 CTGTTTTTTC CCACTAAGT TACACTAACA TCTACCAAG ATAGACTGTT TTTTACAGT 1860
 CAATTTTCA TCACTAACA TATATTAATA CTTTGTAACT TCTTTGCTAT GGCCTTTGTT 1920
 ATCACACCA AACTATGCA TTGTACATG GTTGTTAAG AAGAAACCGT ATTTTCCAT 1980
 GATTAATCAC TGTTTGAAT ATTTGTTCA TGGTATGAT GAAATGTAAA AGCATATTA 2040
 ACACATGGC TGCTAGTTAA CAATTGGAAT AACTTTATTC TGCAATCAT TTAAGAATGA 2100
 ACAGGCGGG CGCTGAGCT CACGCTGTA ATCCAGCAC TTTGGGAGGC TGAGGCGGGC 2160
 AGATCACCTG AGGTGAGGAG TTGGAGACCA GCCTGACCA CATGGACAA CCCGCTCTCT 2220
 ACTAAAAATA CAAAATTGGC AGGTGTGTT GGCACATGCC TATAATCCA GCTACTTGG 2280
 AGGCTAAGC AGGAGATCG CTTGAACCCG GAGGCGGAG GTTGCAGTGA GCCAGATCG 2340
 CACCATGCA CTCTGCTG GCAACAAGA GTGAACTCC ATCTCAAAA A

Seq ID NO: 116 Protein sequence
Protein Accession #: NP_037389

1 11 21 31 41 51
 MALIKAPKRI FGDNFDPDFI KQRRAGLNEF IQNLVRYPEL YNEFDVRAFL QMDSPKHQSD 60
 PSEDEDERSS QKLHSTSQNI NLGSPGNPBA KPTDFDLKV IGKGSFGKVL LAKRKLQKGF 120
 YAVKVLQKKI VLNRKQKHI MAERNVLLN VKHPFLVGLH YSFQITEKLY FVLDFVNGGE 180
 LFFHLQREBS FPEERARFYA ABTASALGYL HSIKIVYRDL KPNILLDVSV GHVVLIDFGL 240
 CKEGIAISDT TTTTCSTPEY LAPEVIRKQP YDNTVDWVCL GAVLYEMLYG LPPFYCRDVA 300
 EMDYNILHKP LSLRPGVSLT AWSILEELLE KDRQNLGAK EDPLEIQNH FFEESLWADL 360
 VQKIPPEFN ENVAGDDIR NEDTAFTEET VPYGVCSVD YSIVNASVLE ADDAFVGFPSY 420
 APPSEDLFL

Seq ID NO: 117 DNA sequence
Nucleic Acid Accession #: NM_004004.1

Coding sequence: 1..681

```

1      11      21      31      41      51
|      |      |      |      |      |
5  ATGGATTGGG GCACGCTGCA GACGATCCTG GGGGGTGTGA ACAAACACTC CACCAGCATT 60
   GGAAGATCTT GGCTCACCCT CCTCTTCATT TTTGCGATTA TGATCCTCGT TGTGGCTGCA 120
   AAGGAGGTGT GGGGAGATGA GCAGGCCGAC TTTGTCTGCA ACACCTTGCA GCCAGGCTGC 180
   AAGAACGTGT GCTACGATCA CTACTTCCCC ATCTCCACCA TCCGGCTATG GGCCCTGCAG 240
   CTGATCTTCG TGTCCAGCCC AGCGCTCCTA GTGGCCATGC ACGTGGCCTA CCGAGACAT 300
10  GAGAGAGAAGA GGAAGTTCAT CAAGGGGGAG ATAAAGAGTG AATTTAAGGA CATCGAGGAG 360
   ATCAAAACCC AGAAGGTCCG CATCGAAGGC TCCCTGTGGT GGACCTACAC AAGCAGCATC 420
   TTCTTCGGGG TCATCTTCGA AGCCGCTTTC ATGTACGTCT TCTATGTTCAT GTACGACGGC 480
   TTCTCCATGC AGCGGCTGGT GAAGTGCAC GCCTGGCCTT GTCCCAACAC TGTGGACTGC 540
   TTTGTGTCCC GGGCCACGGA GAAGACTGTC TTCACAGTGT TCATGATTGC AGTGTCTGGA 600
15  ATTTGCATCC TGCTGAATGT CACTGAATTG TGTATTGTG TAATTAGATA TTGTCTGGG 660
   AAGTCAAAA AGCCAGTTTA A

```

Seq ID NO: 118 Protein sequence
 Protein Accession #: NP_003995.1

```

1      11      21      31      41      51
|      |      |      |      |      |
20 MDWGTIQTIL GGVNKHSTSI GKWLTVLFY FRIMILVVA KEVWGDEQAD FVCNTLOPGC 60
   KNVCYDHYFF ISHRLWALQ LIPVSSPALL VAMBVAYRKH EKRRKFNGE IKSEFKDIEE 120
25 IKTKVRIBG SLHWYTYSSI FFRVFEAAF MYFVYMYDG FSMQLVKCN ANPCPNTVDC 180
   FVSRTETIV FTVFMLAVSG ICILNVTEL CYLLIRYCSG KSKKPV

```

Seq ID NO: 119 DNA sequence
 Nucleic Acid Accession #: XM_061091.1
 Coding sequence: 1..2481

```

1      11      21      31      41      51
|      |      |      |      |      |
35 ATGCCAAATA CTTCAGGAAC AACCAGGATT GAAATTTGGC TTCTCCAAGA GCGGCCCGGG 60
   CACCGAGCGC TGGTGCOCGC TCTCCTTCCG GTGAGTCCCA GCCCGAGATT GGCCTGCGG 120
   CCGGGTACC CCGAGTGCCT GGCCTGCGAT GACCGATTCA CGCTCCGAT GATTGGAGGT 180
   CAGATGCATG GTGAGAGGAT AGATCTCTGG AGCCTTGGTG TTCTTTGCTA TGAATTTTTA 240
   GTTGGGAAGC CTCCTTTTGA GGCAGACGAA GTCCATGTAA GCAAGAAAC CATCGGAG 300
40  ATTTGAGCTG CCAGCAAAAT GATGTGGTGC TCGGCTGCAG TGGACATCAT GTTCTGTGTA 360
   GATGGGTCTA ACAGCGTCGG GAAAGGGAGC TTGAAAGGT CCAAGCATT TGCCATCACA 420
   GTCTGTGAGG GTCTGGACAT CAGCCCGAG AGGCTCAGAG TGGGAGCATT CCAGTTCACT 480
   TCCACTCCTC ATCTGGAATT CCLCTTGGAT TCATTTTCAA CCCACAGGA AGTGAAGGCA 540
   AGATCAAGA GGAATGTTTT CAAAGGAGCG CACAGGAGA CGAACTTGC TCTGAATAC 600
45  CTTCTGCACA GAGGGTTGCC TGGAGGCAGA AATGCTCTG TGCCCGAT CTTCTATC 660
   GTCACTGATG GGAAGTCCCA GGGGGATGTG GCACTGCCAT CCAAGCAGCT GAAGGAAGG 720
   GGTGTCACTG TGTTTGCTGT GGGGGTCAAG TTTCCAGGT GGGAGGAGCT GCATGCACT 780
   GCCAGCGAGC CTAGAGGACA GCACGTCTCT TTGGCTBAGC AGGTGGAGGA TGCCACCAAC 840
   GGCCCTCTCA GCACCTCAG CAGCTCGGCC ATCTGCTCCA GCGCCACGCC AGCTGGGAGC 900
50  CCGAGCTTGG TCTTCATGGA GCGGTTAATG GGCATCTCTC TGATAGGCC CTTGTACTCG 960
   CAGCCCTGCC AGAATGGAGG CACATGTGTT CCAGAAGGAC TGGACGGCTA CCAGTGCTC 1020
   TGCCCGCTGG CCTTTGGAGG GAGGCTAAC TGTGCCCTGA AGCTGAGCCT GGAATGCAAG 1080
   GTCGACCTCC TCTTCTCTCT GGCAGCTCT CCGGGCACCA CTCTGGAGCG CTTCTCTGCG 1140
   GCCAAGTCT TCGTGAAGCG GTTTGTGCGG GCGGTGCTGA GCGAGGACTC TCGGGCCGA 1200
55  GTGGGTGTGG CCACATACAG CAGGGAGCTG CTGGTGGCGG TGCCCTGTGG GAGTACCAG 1260
   GATGTGCTTG ACTGTGCTG GAGCCTCGAT GGCATTCCTT TCCGTGGTGG CCCCACTCG 1320
   ACCGGCAGTG GTCTGCGGCA GCGCGCAGAG CGTGGCTTGG GAGAGCGCAC CAGGACAGCG 1380
   CAGGACCGGC CAGCTAGAGT GGTGGTTTGT CTCACAGT CACACTCCGA GGATGAGGT 1440
   GCGGGCCCGC CGCTCACGC AAGGGCGCGA GAGCTGCTCC TGTGGGTGT AGGCACTGAG 1500
60  GCGGTGCGGG CAGAGCTGGA GGAGATCACA GGCAGCCCA AGCATGTGAT GGTCTACTCG 1560
   GATCCTCAGG ATCTGTTCAA CCAATCCCT GAGCTGCAGG GGAAGCTGTG CAGCCGCGAG 1620
   CGGCCAGGGT GCGGACACA AGCCCTGAGC CTCTCTCTCA TGTGGACAC CTCGCTCTCA 1680
   GTAGGGCCCG AGAATTTTGC TCAGATGAGC AGCTTTGTGA GAAGCTGTGC CCTCCAGTTT 1740
   GAGGTGAACC CTGACCTGAC ACAGGTCCGC CTGGTGGTGT ATGGCAGCCA GGTGCAGACT 1800
65  GCCTTCGGGC TGGACACCAA ACCCACCAG GCTGCGATGC TGGGGCCAT TAGCCAGGCC 1860
   CCTTACCTAG GTGGGGTGGG CTCAGCCGCG ACCGCCCTGC TGACATCTTA TGACAAAGTG 1920
   ATGACCTTCC AGAGGGGTGC CCGGCTGGT GTCCCAAAG CTGTGGTGGT GCTCACAGGC 1980
   GGGAGAGGCG CAGAGGATGC AGCCGTCTCT GCCCAGAAAG TGAGGAACAA TGGCATCTCT 2040
70  GTCTGTGTGG TGGCGGTGGG GCCTGTCTCA AGTGAGGCTC TGGGAGGCT TGCAGGTCCC 2100
   CGGATTTCCT TGAATCAAGT GGCAGCTTAC GCCGACCTGC GGTACCAACA GGCAGTCTC 2160
   ATTGAGTGGC TGTGTGAGGA AGCCAAGCAG CCAGTCAACC TCTGCAAAAC CAGCCCGTGC 2220
   ATGAATGAGG GCGAGCTGGT CTTGCAAAAT GGGAGCTACC GCTGCAAGTG TGGGATGGC 2280
   TGGAGGGGCC CCACTGCGA GAACCGTAG TGGAGCTCT GCTCTGTATG TGTGAGCCAG 2340
75  GGTATGATTC TTGAGACGCC CCTGAGGCC ATGGCTCCCG TGCAGGAGG CAGCAGCCCT 2400
   ACCCTCCCA GCAACTCAG AGAAGGCTTG GGCCTGAAA TGGTGCCTAC CTTCTGGAAT 2460
   GTCTGTGCC CAGTCTCTA G

```

Seq ID NO: 120 Protein sequence
 Protein Accession #: XP_061091.1

```

1      11      21      31      41      51
|      |      |      |      |      |
80 MNVTSGITRI SIWLLQEPFG HRALVAALLS VSPSPRLALA PGYFPVPAAD DRFTLEMIIG 60
   QMBEKVDLW SLGLVLYEFL VGRPPPEANE VHVSKETIGK ISAASKMMWC SAAVDIMFL 120

```

5 DGSNSVGKGS FERSKHEFAIT VCDGLDISPE RVRVGAFQFS STPHLEPFLD SPSTQQEVKA 180
 RIKRMVFKGG RTETELALKY LLHRGLFQGR NASVPQILII VTDGKSQGDV ALPSKQLKER 240
 GVTTFVAVGVR PPRWEELHAL ASEERGGHVL LAEQVEDATN GLFSTLSSSA ICSSATPAGS 300
 PELVPMERLM GISLIGPCDS QPCQNGGTCV PBGLDGYQCL CPLAFGGGAN CALKLSECR 360
 VDLLFLDSS AGTTLDGFLR AKVPVKRFVR AVLSSEDSRAR VGVATYSREL LVAVPVGEYQ 420
 DVEDLVWSLD GIPFRGGPTL TGSALRQAAE RGFGSATRTG QDRPRRVVVL LTESHSEDEV 480
 AGPAKHARAR ELLLLGVGSE AVRAELEETI GSPKHVMVYS DPQDLFNQIP ELQGLKCSRQ 540
 RPGCRTOALD LVFMLDTSAS VGPENFAQM Q SFVRSCALQF EVNPDVTQVG LVVYGSQVQT 600
 10 AFGLDTKPTR AAMLRASQA PYLGGVGSAG TALLHIYDKV MTVQRGAREG VPKAVVVLTG 660
 GRGADAAVF AQKLNNNGIS VLVVGVGPVL SEGLRRLAGP RDSLIHVAAY ADLRYHODVL 720
 IENLCEGAHQ PVNLCKPSPC MNEGSCVLQN GSYRCKCRDG WEGPHCENRE WSSCSVCVSQ 780
 GWILETFLRH MAPVQEGSSR TPPSNYREGL GTEMVPTFWN VCAPGP

15 Seq ID NO: 121 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..2424

20 1 11 21 31 41 51
 ATGCCCCCTT TCCTGTGTCT GGAGGCCGTC TGTGTTTTCC TGTTTTCCAG AGTCCCCCCA 60
 TCCTCTCCCTC TCCAGGAAGT CCATGTAAGC AAGAAACCA TCGGGGAAGAT TTCAGCTGCC 120
 AGCAAAATGA TGTGTGCTC GGTGCGAGTG GACATCATGT TTCTGTTAGA TGGGTCTAAC 180
 AGCGTCGGGA AAGGGAGCTT TGAAGGTCC AAGCATTGT CCATCAGAGT CTGTGACGGT 240
 25 CTGGACATCA GCCCCGAGG GGTGAGAGTG GAGCATTCC AGTTCAGTTC CACTCCTCAT 300
 CTGGAATTCC CTTGGATTTC ATTTTCRACC CAACAGGAAG TGAAGGCAAG AATCAGAGAG 360
 ATGGTTTTC AAGGAGGGCG CACGGAGAGC GAACCTGTCT TGAATAACCT TCTGCACAGA 420
 GGGTGCTCT GAGGCAAGAA TGTCTGTGTG CCCAGATCC TCATCATCGT CACTGATGG 480
 AAGTCCAGG GGAATGTGGC ACTGCCATCC AAGCAGCTGA AGGAAAGGGG TGTCACTGTG 540
 TTTGCTGTGG GGTTCAGGTT TCCCAGGTGG GAGGAGCTGC ATGCACTGGC CAGCAGGCT 600
 30 AGAGGGCAGC AGTGTCTGTT GGTGAGGATG GTGAGGATG CCACCAACGG CCTCTTCAGC 660
 ACCCTCAGCA GCTCGGCCAT CTGCTCCAGC GCCACGCCAG ACTGCAAGGT CGAGGCTCAC 720
 CCTGTGAGC ACAGGACGCT GGAGATGGTC CGGGAGTGG CTGGCAATGC CCATGCTGG 780
 AGAGGATCCB GCGGACCCCT TGCGGTGCTG GCTGCACACT GTCCCTTCTA CAGCTGGAAG 840
 AGAGTGTTC TAACCCACCC TGCCACCTGC TACAGGACCA CTTGCCAGG CCCCCTGTAC 900
 35 TCGCAGCCCT CCAGAAATGG AGGCACATGT GTTCCAGAA GACTGGACGG CTACCAAGTC 960
 CTCTGCCCGC TGCCCTTTGG AGGGGAGGCT AACTGTGCC TGAAGCTGAG CTTGGAATGC 1020
 AGGTGCGACC TCTCTCTCT GCTGGACAGC TCTGCGGCA CCACTCTGGA CCGCTTCTTG 1080
 CCGGCCAAG TCTGTGTAA GCGGTTGTG CCGGCCGTGC TGAGCGAGGA CTCTCGGGCC 1140
 40 CGAGTGGGTG TGCCACATA CAGCAGGAG CTGCTGTGG CCGTGCCTGT GGGCGAGTAC 1200
 CAGGATGTC CTGACCTGGT CTGAGCCTC GATGGCATTC CCTTCGTGG TGGCCCCACC 1260
 CTGACGCGCA GTCCCTTGG GCAAGCGGCA GAGCGTGGCT TCGGAGGCGC CACCAAGACA 1320
 GGCAGGACC GCGCAGGTAG AGTGGTGGT TGTCTCACTG AGTCAACTC CAGGATGAG 1380
 GTTGGCGGCC CAGCGGTCA CCGAGGGCG CAGAGAGCTG TCTGCGGGG TGTAGGCAAT 1440
 45 GAGGCGGTGC GGCAGAGCT GAGGAGATC ACAGGACGCC CAAGCATGT GATGGTCTAC 1500
 TCGGATCCTC AGGATCTGTT CRACCAATC CCGAGCTGC AGGGGAGCT GTGAGCCCG 1560
 CAGCGGCGAG GGTCCCGGAC ACAAGCCCTG GACCTCGTCT TCAATGTTGA CACTCTGCC 1620
 TCAGTAGGCG CCGAATATT TGTCTAGAT CAGAGCTTTC TGAGAAGCTG TGCCCTCCAG 1680
 TTTGAGGTGA ACCCTGACCT GACACAGCT GGCCTGTGG TGTATGGCAG CAGGTGCGAG 1740
 50 ACTGCTCTGG GGTGGACAC CAACCCACC CCGGCTGCGA TGCTGCGGC CATTAGCCAG 1800
 GCCCCCTACC TAGGTGGGT GGGCTCAGCC GGCACCGCCC TGCTGCACAT CTATGACAAA 1860
 GTGATGACCG TCCAGAGGGG TGCCCGGCTT GGTGTCCCA AAGCTGTGGT GGTGCTCACA 1920
 GCGCGGAGAG GCGCAGGTA TGCAAGCGTT CCGTCCCGA AGCTGAGGAA CAATGGCATT 1980
 TCTGTCTGG TCGTGGCGT GGGGCTGTG CTAAGTGAG GTCTGCGGAG GCTTGAGGT 2040
 55 CCCCCGGATT CCTGTATCCA CGTGGCAGCT TACGCGGACC TGCGGTACCA CAGGACCTG 2100
 CTCAATTGAG GGTGTGTGG AGAAGCCAG CAGCCAGTCA ACCTCTGCAA ACCAGCCCG 2160
 TGCATGATG AGGCGAGCTG CGTCTGCGAG AATGGAGCT ACCGCTGCAA GTGTGCGGAT 2220
 GCGTGGAGG GCGCCCACTG CAGAACCGT GAGTGGAGCT CTTGCTCTGT ATGTGTGAGC 2280
 CAGGATGGA TTCTTGAGAC GCGCCGAGG CACATGGCT CCGTGCAGGA GGGCAGCAGC 2340
 60 CGTACCCCTC CCAGCACTA CAGAGAAGGC CTGGGCACTG AATGGTGCC TACCTTCTGG 2400
 AATGTCTGTG CCCAGGTCC TTAG

Seq ID NO: 122 Protein sequence
 Protein Accession #: Eos sequence

65 1 11 21 31 41 51
 MPFFLLLEAV CVFLFSRVPP SLPLQEVHVS KETIGKISAA SKMMWCSAAV DIMFLDGSN 60
 SVGKGSFERS KFAITVCDG LDISPERVRV GAFQFSSTPH LEFPLDSFST QQSVKARIKR 120
 70 MVFKGGRTET ELALKYLLER GLPGGRNASV PQLLIIVTDG KSQGDVALPS KQLKRGVTV 180
 FAVGVRFPRN EELHALASEP RGQHVLLAEQ VEDATNGLFS TLSSSAICSS ATEDCRVEAH 240
 BCEHRTLEMV REFAGNAPCM RGSRTLAIVL AAHCPEYSWK RVFLTHPATC YRTTCPGPFD 300
 SQPCQNGGTC VEGLDGYQC LCPAFQGEA NCALKLSLEC RVDLLFLDSS SAGTTLDGFL 360
 RAKVPVKRFV RAVLSEDSRA RVGVATYSRE LVAVPVGEY QVFDLVWSL DGIPFRGGPT 420
 LTGSALRQAA ERGFGSATRT QDRPRRVVVL LTESHSEDEV VAGPAKHARA RELLLGVGS 480
 75 EAVRAELESI TSPKHVMVY SDPQDLFNQI PELQGLKCSR QRPGRCTQAL DLVFMLDISA 540
 SVGPENFAQM QSFVRSCALQ FEVNPDTVQV GLVYGSQVQ TAFGLDTKPT RAAMLRASQ 600
 APFLGGVGS A SEGLRRLAG PRDSLIHVA YADLRYHODV LLENLCEGAH QFVNLCKPSP 660
 80 CMWEGSCVLQ NGSYRCKCRD GWEGPHCENR WSSCSVCVS QWILETFLRH MAPVQEGSS 720
 RTPSNYREGL LGTEMVPTFW NVCAFGP

Seq ID NO: 123 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 89..2356

	1	11	21	31	41	51	
5	GCCCCCTGGC	CCGAGCCGCG	CCCGGGCTCG	TGAGTAGAGC	CGCCCGGGCA	CCGAGCGCTG	60
	GTGCGCGCTC	TCCTTCGGTT	ATATCAACAT	GCCCCCTTTC	CTGTTGCTGG	AAGCCGCTCG	120
	TGTTTTCCCTG	TTTTCCAGAG	TGCCCCCATC	TCTCCCTCTC	CAGGAAGTCC	ATGTAAGCAA	180
	AGAAACCATC	GGGAAGATTT	CAGCTGCCAG	CAAAATGATG	TGGTGTCTGG	CTGCAGTGGG	240
	CATCATGTTT	CTGTAGATG	GGTCTAACAG	CGTCGGGAAA	GGGAGCTTTG	AAAGGTCCAA	300
10	GCACCTTGCC	ATCAGAGTCT	GTGACGGTCT	GGACATCAGC	CCCGAGAGGG	TCAGAGTGGG	360
	AGCAITCCAG	TTCCATTTCA	CTCCTCATCT	GGATTTCCTC	TTGGATTTCAT	TTTCAACCCA	420
	ACAGGAAGTG	AAGGCAAGAA	TCAAGAGGAT	GGTTTTCAAA	GGAGGGCGCA	CGGAGACGGA	480
	ACTTGCTCTG	AAATACCTTC	TGCACAGAGG	GTGCGCTGGA	GGCAGAAATG	CTTCTGTGCC	540
	CCAGATCCCTC	ATCATCGTCA	CTGATGGGAA	GTCCCGAGGG	GATGTGGCAC	TGCCATCCAA	600
15	GCAGCTGAAG	GAAAGGGGTG	TCACGTGTGT	TGCTGTGGGG	GTGAGGTTTC	CCAGGTGGGA	660
	GGAGCTGCAT	GCACGTGCCA	GCAGGCTTAG	ABGGCAGCAC	GTGCTGTTCG	CTGAGCAGGT	720
	GGAGGATGCC	ACCAACGGCC	TCCTCAGCAC	CCTCAGCAGC	TGCGCATCTC	GCTCCAGCGC	780
	CAGCCGACAG	CAGCGGCTCG	AGGCTCACCC	CTGTGAGCAC	AGGACGCTGG	AGATGGTCCG	840
	GGAGTTCGCT	GGCAATGCCC	CATGCTGGAG	AGGATCGCGG	CGGACCTTTC	CGGTGCTGGC	900
20	TGCACACTGT	CCCTTCTACA	GCTGGAGAGB	AGTGTTCCTA	ACCCACCTTC	CCACTGCTTA	960
	CAGGAACACC	TGCCCAGGCC	CCTGTGACTC	GCAGCCCTGC	CAGAAATGGAG	GCACATGTGT	1020
	TCAGAAAGGA	CTGGACGGCT	ACCATGTCCT	CTGCCCGCTG	GCCTTTGGAG	GGGAGGCTAA	1080
	CTGTGCCCTG	AAGCTTGAGC	TGGAATGCAG	GGTCGACCTC	CTCTTCTGTC	TGGACAGCTC	1140
	TGCGGGCACC	ACTCTGGAGC	GCTTCTGCTG	GGCCAAAGTC	TTCTGTGAGC	GGTTTGTGCG	1200
25	GGCGTGTCTG	AGCGAGGACT	CTCGGGGCGG	AGTGGGTGTG	GCCACATACA	GCAGGGAGCT	1260
	CGTGTGGCGG	GTGCGGTGTG	GGAGTACCA	GGATGTCCCT	GACCTGGTCT	GGAGCCTCGA	1320
	TGGCATTCCT	TTCCGTGGTG	GGCCCAACCT	GACGGGCGAG	GCCTTGCCTG	AGGCGGCAGA	1380
	GGTGGCTTTC	GGGAGGCGCA	CCAGGACAGG	CCAGGACCGG	CCAGTATAGG	TGGTGGTTTT	1440
	GCTCACTGAG	TCACACTCCG	AGGATGAGGT	TGCGGGCCCA	GCCTGTGCTG	CAAGGGGCGG	1500
30	AGAGCTGCTC	CTGCTGGGTG	TAGGCACTGA	GGCCGTGCGG	GCAGAGCTGG	AGGAGATCAC	1560
	AGGCAGCCCA	AGCATCTGTA	TGCTCTACTC	GGATCCTCAG	GATCTGTTCA	ACCAATCCCT	1620
	TGAGCTGCAG	GGGAGCTGT	GCAGCCGCGA	CGCGCCAGGG	TGCGGACAC	AAGCCCTGGA	1680
	CCTGCTCTTC	ATGTTGGACA	CCTCTGCTTC	AGTAGGGCCC	GAGAAITTTG	CTCAGATGCA	1740
	BAGCTTTGTG	AGAGCTGTG	CCCTCCAGTT	TGAGGTGAAC	CCTGACGTGA	CACAGGTCCG	1800
35	CCTGTGGTGT	TATGGCAGCC	AGGTGCAGAC	TGCTTTCGGG	CTGGACACCA	AACCCACCGG	1860
	GGCTGCGATG	TGCGGGGCA	TTAGCCAGGC	CCCTACCTTA	GGTGGGGTGG	GCTCAGCCGG	1920
	CACCGCCCTG	CTGCATCACT	ATGACAAAGT	GATGACCTGC	CAGAGGGGTG	CCCGGCCCTG	1980
	TGTCGCCAAA	CTGTGGGTGG	TGCTCACAGG	CGGGAGAGGC	GCAGAGGATG	CAGCCCTTCC	2040
	TGCCCCAGAG	CTGAGGAACA	ATGGCATCTC	TGCTCTGGTC	GTGGGCTGTG	GGCTGTCTCT	2100
40	AAGTGAGGGT	CTGCGGAGGC	TTGCAGGTCC	CCGGGATTC	CTGATCCACG	TGGCAGCTTA	2160
	CGCGGACCTG	CGGTACCACC	AGGACGTGCT	CATTGAGTGG	CTGTGTGGAG	AGGCCAAGCA	2220
	GGCACTCAAC	CTCTGCAAAC	CCAGCCCGTG	CATGAATGAG	GGCAGCTGCG	TCTGTGAGAA	2280
	TGGGAGCTAC	CGCTGCAAGT	GTGGGATGCG	CTGGGAGGCG	CCCACTGCGG	AGAACCGATT	2340
	CTTGAGACBC	CCCTGAGGCA	CATGGCTCCC	GTGCAGGAGG	GCAGCAGCCG	TACCCCTCCC	2400
45	AGCACTACA	GAGAAGGCGT	GGGCACTGAA	ATGGTGCTTA	CTTCTGAGAA	TGCTGTGCTC	2460
	CCAGCTCTCT	AGATATGCTG	CTTCCGCGCG	TGGCCAGGAC	CACATATCTC	ACTGAGGGAG	2520
	GAGGATGTCC	CACTGCGAGC	CATGCTGCTT	AGAGACAGAA	AAGCAGCTGA	TGTCACCCAC	2580
	AAACGATGTT	GTGTGAAAGT	TTTGATGTGT	AAGTAAATAC	CCACTTTCCT	TACCTGCTGT	2640
	GGCTGTGTGA	GGCTATGTCA	TCGTGCACTC	TTCCCTTGAG	GATAAACAAG	GGGTCTGTGA	2700
50	GACTTAAATT	TAGCGGCTCG	AGGTTCTCTT	GCACACATATC	AATGCTCGCC	AGAATGTTGT	2760
	TGACACAGTA	ATGCCAGCA	GAGGCCTTTA	CTAGAGCATC	CTTTGACGCG		

Seq ID NO: 124 Protein sequence
Protein Accession #: Bos sequence

55	1	11	21	31	41	51	
	MPFFLLLEAV	CVFLFSRVPP	SLPLQEVHVS	KETIGKISAA	SKMMWCSAAV	DIMFLLDGSH	60
	SVGKGSFERS	KHFAITVCDG	LDISPERVRV	GAPQFSSTFH	LEFFLDSPST	QSEVKARIKR	120
60	MVFRGGRTET	ELALKYLLER	GLPGGRNASV	PQILIVITDG	KSQGDVALPS	KQLKRGVTV	180
	PAVGVRFRPM	RELHALASEP	RQDEVLLAEQ	VEDATNGLEFS	TLSSSAICSS	ATPDCRVEAH	240
	PCEHRTLEMV	REFAGNAPCH	RGRRTLAVL	AAHCPFFYSWK	RVFLTHPATC	YRTTCFPGCD	300
	SQPCQNGGTC	VPEGLDGYQC	LCFLAFGGEA	NCAKLKSLSC	RVDLLFLDLS	SAGTTLDGFL	360
	RAKVVRKRFV	RAVLESDSRA	RVGVATYSRE	LLVAVFVGEY	QDVEDLVMSL	DGIFPRGGPT	420
65	LTGSALRQAA	ERFGGSATRT	GQDRPRRVVV	LLTESHSEDE	VAGPARHARA	RELLLLGVGS	480
	EAVRAELEEI	TGSEPKHVMVY	SDPDLEFNQI	PBLQKLCNSR	QRPGCRTQAL	DLVFMIDTSA	540
	SVGFENFAQM	QSFRVSCALQ	FEVNPDTVQV	GLVVYGSQVQ	TAFGLDTKPT	RAAMLRAISQ	600
	APYLGGVGS	GTALLHIYDK	VMTVQRGARP	GVPKAVVVL	GGRGABDAAV	PAQKLRLNNGI	660
	SVLVVGVGPFV	LSEGLRLRLAG	PRDELHVA	YADLRYHQDV	LIEMLCGEAK	QPVNLCCKPSP	720
70	CMNBSGCVLQ	NGSYRCKCRD	GWEGPHCENE	FLRRP			

Seq ID NO: 125 DNA sequence
Nucleic Acid Accession #: NM_031942.1
Coding sequence: 145..1260

75	1	11	21	31	41	51	
	CCGAGGCCCC	GGCCCTCCGG	GGCCGGGTCTG	GGCGGCCGAG	CCTGCGAGCC	GGCTGTCTGC	60
	TGCTCTCTCT	GCTGTGGGAC	CGCTGACCGC	GGCGCTGCTC	CGCTCTCCCC	GCTCCAGCG	120
	CGGATCTGGG	CACCCGCCAC	CAGCATGGAC	GCTCGCCGCG	TGCGGCAGAA	AGATCTCAGA	180
80	GTAAAGAAGA	ACTTAAAGAA	ATTGAGATAT	GTGAAGTTGA	TTTCCATGGA	AACCTCTGCA	240
	TCCCTCTGAG	ACAGTTGTGA	CAGCTTTGCT	TCTGATAATT	TTGCAAAACAC	GAGGCTGCG	300
	TTCAGTCTGG	AAGGCTGTAG	GACCCGCGAG	CAGTGCAGGC	ACTCTGGACC	TCTCAGGGTG	360
	GGGATGAAGT	TTCCAGCGCG	GAGTACCAAG	GGAGCAACCA	ACAAAAGAC	AGAGTCCCGC	420
	CAGCCCTCAG	AGAATTCTGT	GACTGATTCC	AACTCCGATT	CAGAAGATGA	AAGTGGGAATG	480

5 AATTTTITGG AGAAAAAGGC TTTAAATATA AAGCAAAACA AAGCAATGCT TGCAAAACTC 540
 ATGTCTGAAT TAGAAAGCTT CCTGGGCTCG TTCCGTGGAA GACATCCCCT CCCAGGCTCC 600
 GACTCACBAAT CAGAGGAGCC GCGAGGCGGT ACATTCCTCG GTGTTGCTTC CAGGAGAAAC 660
 CCTGAACGGA GAGCTCGTCC TCTTACCAGG TCAGGTGCCC GGATCCTCGG GTCCCTTGAC 720
 GCTCTACCCA TGGAGGAGGA GGAAGGAGAG GATAAGTACA TGTGGGTGAG AAAGAGGAAG 780
 ACCGTGGATG GCTACATGAA TGAAGATGAC CTGCCAGAA GCGGTGCTCT CAGATCATCC 840
 GTGACCCCTC CCGATATAAT TCGCCCACTG GAAGAAATTA CAGAGGAGGA GTTGGAGAAC 900
 GTCTGCAGCA ATTCTCGAGA GAAGATATAT AACCGITCAC TGGGCTCTAC TTGTTCATCA 960
 TGCCGTCAGA AGACTATTGA TACCAAAACA AACTGCAGAA ACCCAGACTG CTGGGCGGTT 1020
 CGAGGCCAGT TCTGTGGCCC CTGCCCTCGA AACCGTATG GTGAAGAGGT CAGGGATGCT 1080
 CTGCTGGATC CGAACTGGCA TTGCCCGCCT TGTGAGGAA TCTGCAACTG CAGTTTCTGC 1140
 CGGCAGCGAG ATGGACGGTG TCGGACTGGG GTCCCTGTGT ATTTAGCCAA ATATCATGGC 1200
 TTTGGGAATG TGCATGCCIA CTTGAAAAGC CTGAACAGG AATTGGAAT GCAGCATAA 1260
 TATCTGGAAG ATTTGCTGCC TGCCCTCTAC TTCTCAATC TTCTGTGTA AAGTTTCCAA 1320
 TTTTTCACCT GAAACCTGAG TTAATAATCT TGATGATCAG CCTGTTTCAT AAGAACTCC 1380
 AATCAAGTTA ATCTTAGCAG ACATGTGTTT CTGGAGCATC ACAGAAAGTA TATTGCTAGT 1440
 TACACITTCG CCTCTGCGAG TTTCTTCTCT GCTCCCAACC CCCATCTCAT AGCATCCCC 1500
 TCAATTTCCA ATGCTCCTCT CCAACCGCTT AGTTTCTGAA TTTCTTTTAA ATTACAGTTT 1560
 TATGAAACGA ATTTTATTT ACTTGGTGT GAAATAGCCC TCATAAAACC TAAGCACTTG 1620
 GAAACACAA ATATGATTTA ACTAATGAGA TCTATTGAAT TTCAGAGAG AGCCTTCTAA 1680
 CTGTGTTACA CAAAACGAG TATGATTAG CACTCATACT AGTTGAAAT TTTAATAGAA 1740
 TCAAGGCACA AAAGCTTAA AACCATGAG AAAAATTAG TAATTATTGC AGATTGATGT 1800
 CTCTCAATCC CAGTATATGC GCTTATGTTA CAGTTGTTG TCACAGTTGA GACTTAATTT 1860
 CTCTTAATTT CTCTGCCCC AGGGTAAAGT GGTGCGTCCA GCTTACACGA TCATAATTCA 1920
 AAGGTTGGTG GGCATGTAA TACTTAATTA AATAATGAT GGAAGAGCTA TCTGGAGATT 1980
 ATGAGTAAGC TGATTGAAAT TTTCAATATA AAACCTTAGT ATAAATGAG TTTGCAAGT 2040
 TTATTTCAAT TCACATGTAA GGTATTGCAA ATAAATCTT GACAAATTT GTATGGAAAC 2100
 TTGATATTA AAACAGTCT GTGCTTCTT GCAGTTCTT GTAAATTTAT AAACAGGCA 2160
 CAGGTTTCAA GTTTAGATT TAAGCACTTT TATAACAATG ATAAATGCTT TTTTGGAGAT 2220
 GTAACITTTA CGAGTTTGT AACTGACAT CTCTGCCAG CTAGTTTCTG GGCAGGTTTC 2280
 CAGGTGTCAG ATTCGCCCTC CTCTTGCAT TAATCAAGGT ATTTGGTACA GGTGGAATCT 2340
 AAGTGTTTGT ATGTCCAAT TACTTGCATA TGTAAACCAT TGCTGTGCCA TTCAATGTTT 2400
 GATGCAATAT TGGACCTTGA ATCGATAAGT GTAAATACAG CTTTGTATCT GTAATGCTTT 2460
 TATACAAAG TTTATTTTAA TAATAAATG TTTGTTCTAA AAAAAAATA

Seq ID NO: 126 Protein sequence
Protein Accession #: NP_114148.1

40 1 11 21 31 41 51
 MDARRVPQKD LRVKKLKKF RYVKLISMET SSSDDSDCDS FASDNFANTR LQSVREGCRT 60
 RSCCRHSGLP RVAMKFPARS TEGATNKKAE SRQPSNSVT DSNDSDEDS GMPLEKRAL 120
 NIKQNKANLA KLMSELESFP GSFGRHPLP GSDSQRRPR RPTFPGVASR NPFERRARPL 180
 45 TRSRRLIGS LDALPMEEES EEDXYMLVRK RKTVDGYMNE DDLPRSRSR SSVTLPHIIR 240
 PVEITTEEL ENVCNSREK IYNRSLGSTC HQCRQKIDT KTNCRNPDWN GVRGQFCGPC 300
 LRNRXGEEVR DALLDENWHC PPCRGICNCS PCRQRDGRCA TGVLVYLAKY HGFQNVHAYL 360
 KSLKQEFEMQ A

Seq ID NO: 127 DNA sequence
Nucleic Acid Accession #: AF305616.1
Coding sequence: 1..863

55 1 11 21 31 41 51
 ATGCACCGCT TGATGGGGGT CAACAGCACC GCGCCGCGCG CCGCCGGGCA GCCCAATGTC 60
 TCCTGCACGT GCAACTGCAA ACCTCTTTG TTCCAGAGCA TGGAGATCAC GGAGCTGGAG 120
 TTGTGTCAGA TCATCATCAT CGTGGTGGTG ATGATGGTGA TGGTGGTGGT GATCAGGTGC 180
 CTGCTGAGCC ACTACAGCT GTCTGCACGG TCCTTCATCA GCGGCGACAG CCAGGGGCGG 240
 60 AGGAGAGAAG ATGCCCTGTC CTCAGAAAGA TGCCGTGGG CCTCGAGAG CACAGTGTCA 300
 GGCAACGGA TCCAGAGCC GAGGTCTAC GCGCCGCTC GCGCCACCGA CCCTGTGGCC 360
 GTGCCGCCCT TCGCCAGCG GAGCGCTTC CACCGCTTC AGCCACCTA TCCGTACCTG 420
 CAGCAGGAGA TCGACCTGCC ACCACCATC TCGCTGTGAG ACCGGGAGGA GCCCCACCC 480
 TACCAGGGCC CCGTGCACCT CCAGCTTCGG GACCCGAGC AGCAGCTGGA ACTGAACCG 540
 65 GAGTGGGTGC GCGCACCCCT AACAGAACCT ATCTTCGACA GTGACCTGAT GGTAGTGTCC 600
 AGGCTGGGCG GCCCTGTGCC CCCAGCAGT AACTCGGGA TCAGCGCCAG GTGCTAGGCG 660
 AGCGCGGGGC GCATGGAGGG GCGCCGCCCT ACCTACAGCG AGGTCAATCG CCACTACCCG 720
 GGGTCTCTCT TCCAGCACA GCAGCAGCT GCGCCGCCCT CCTTGTCTGA GGGGACCCG 780
 70 CTCCACACA CACACATCG GCCCTAGAG AGCGCAGCA TCTGGAGCAA AGAGAAGGAT 840
 AACAGGAAG GACACCTCT CTAG

Seq ID NO: 128 Protein sequence
Protein Accession #: AAL09357.1

75 1 11 21 31 41 51
 MRLMGNVNST AAAAAGQENV SCTCNKRSL FQSMETILE FVQIIIVVV MVMVNVVITC 60
 LLSHYKLSAR SFISKESQGR RREDALSSSG CLWPSESTVS GNGIPEQVY APPRPTDRLA 120
 VPPFAQRRER HRPQETPYPL QHEIDLPPYT SLSDGEEPPP YQGPCTQLR DPEQQLNLN 180
 80 ESVRAPPNRT IFDSLDMSA RLGGPCFPSS NSGISATCYG SGRMEGPPF TYSEVIGHYP 240
 GGSFQHQQS GPPSLLEGTR LHRTHIAPLE SAAIWSKED KQKQHP

Seq ID NO: 129 DNA sequence
Nucleic Acid Accession #: NM_004952.1

Coding sequence: 1..718

5
10
15

1	11	21	31	41	51	
ATGGGCGGCG	CTCCGCTGCT	GCTGCTGCTG	CTGCTGGTGC	CGGTGCCGCT	GCTGCCGCTG	60
CTGGGCCAAG	GGCCCGGAGG	GGCGCTGGGA	AACCGGCATG	CGGTGTACTG	GAACAGCTCC	120
AACCGACACC	TGCGCGGAGA	GGGCTACACC	GTGCAGGTGA	ACGTGAACGA	CTATCTGGAT	180
ATTACTGCC	CGCACTACAA	CAGCTCGGGG	GTGGGCCCCG	GGCGGGGACC	GGGGCCCCGA	240
GGCGGGGAG	AGCAGTACGT	GCTGTACATG	GTGAGCCGCA	ACGGCTACCG	CACCTGCAAC	300
GCCAGCCAGG	GCTTCAAGCG	CTGGGAGTGC	AACCGGCCGC	ACGCCCCGCA	CAGCCCCATC	360
AAGTTCTCGG	AGAAGTTCCA	GCGCTACAGC	GCCTTCTCTC	TGGGCTACGA	GTTCACGCC	420
GGCCACGAGT	ACTACTACAT	CTCCACGCC	ACTCACAACC	TGCCTGGAA	GTGTCTGAGG	480
ATGAAGGTGT	TGCTCTGCTG	CGCTCCACA	TGCACTCCG	GGGAGAACCC	GGTCCCCACT	540
CTCCCCAGT	TCACCATGGG	CCCCAATGTG	AAGATCAACG	TGCTGGAGA	CTTTGAGGGA	600
GAGAACCCTC	AGGTGCCCAA	GCTTGAGAAG	AGCATCAGCG	GGACCAAGCC	CAACCGGAA	660
CACCTGCCCC	TGGCGGTGGG	CATCGCCTTC	TTCTCATGA	GTTCTTGGC	CTCCTAG	

Seq ID NO: 130 Protein sequence
Protein Accession #: NP_004943.1

20
25

1	11	21	31	41	51	
MAAAPLLLLL	LLVPVPLPL	LAQPGGALG	NRHAVYVMS	NQHLRREGYT	VQVNVNDYLD	60
IYCPHYNSSG	VCPGAGGPG	GGABQVLYM	VSRNGYRTCN	ASQGFKNWEC	NRPHAPHSPI	120
KFSEKFORYS	AFSLQVFEHA	GHEYYVISTP	THNLHWKCLR	MKVYVCCAST	SHSGKGVPT	180
LPQFTMGPNV	KINVLDFEG	ENPQVPLBK	SISGTSFKRE	HLPLAVGIAP	FLMTFLAS	

Seq ID NO: 131 DNA sequence
Nucleic Acid Accession #: NM_012445.1
Coding sequence: 276..1271

30
35
40
45
50
55
60
65

1	11	21	31	41	51	
GCACGAGGGA	AGAGGGTGT	CCGACCCGGG	GAAGGTGCTT	GGGCAAGGCG	AGITGGGAAA	60
GGCGGAGCCC	CGCGCCGCC	CGCAGCCCTT	TCTCTCTCTT	TCTCCACGT	CCTATCTGCC	120
TCTCGCTGGA	GGCCAGGCCG	TGCAGCATCG	AAGACAGGAG	GAAGTGGAGC	CTCATTTGCC	180
GGCCCGGGGC	GGCGGCTTCG	GGCTTAAATA	GGAGCTCCCG	GCTCTGGCTG	GGACCCGACC	240
GCTGCGGGCC	GGCTCCCGCC	TGCTCTCTCC	GGGTGATGGA	AAACCCAGC	CGGCGCGCCG	300
CGCTGGGCAA	GGCGCTCTGC	GCTCTCTCTC	TGGCCACTCT	CGGCGCGCC	GGCCAGCCTC	360
TTGGGGGAGA	GTCCATCTGT	TCCGCCAGAG	CCCGGCCCAA	ATACAGCATC	ACCTTCACGG	420
GCAAGTGGAG	CCAGACGGCC	TTCCCAAGC	AGTACCCCTT	GTTCCGCCCC	CCTGCGCAGT	480
GGTCTTCGCT	GCTGGGGGCC	GGCATAGCT	CXGACTACAG	CATGTGGAGG	AAGAACCACT	540
ACGTAGTAA	CGGCTCGGCG	GACTTTGCGG	AGCGCGGCGA	GGCTTGGGCG	CTGATGAAGG	600
AGATGAGAG	GGCGGGGGAG	GGCTGCGAGA	GGTGCACGCG	GGTGTCTTGG	GGCGCGGCGG	660
TCCCGAGCGG	CACCGGGCAG	AGCTGCGCGG	AGCTGGAGGT	GCACGCGCAG	CACCTCAGTG	720
TCTCGTTTGT	GGTGGCATC	GTGCCAGGCC	CGACTGGTTT	CGTGGGCGTG	GACAGCCTGG	780
ACCTGTGCGA	CGGGGACCGT	TGGCGGGAAC	AGCGGGCGCT	GGACCTGTAC	CCCTACGACG	840
CGGGGACGGA	CAGCGGCTTC	ACCTTCTCTT	CCCCCAACTT	CGCCACCAAT	CCGCGAGGCA	900
CGGTGACGGA	GATAACGCTC	TCTCTCTCCA	GCCACCGGCG	CAACTCTCTT	TACTACCCGC	960
GGCTGAAGGC	CCTGCTCTCC	ATCGCCAGGG	TGACACTGGT	GGGCTGCGGA	CAGAGCCCCA	1020
GGGCTTCAT	CCCTCCGCGC	CCAGCTCTGC	CCAGCAGGGA	CAATGAGATT	GTAGACAGCG	1080
CTTCAGTTCC	AGAAAGCGCG	CTGGACTGCG	AGGTCTCTCT	GTGGTCTGTC	TGGGACTCTG	1140
GGGAGGCGCA	CTGTGGGAGG	CTGGGACCA	AGAGCAGGAC	TGGTACGTC	CGGGTCCAGC	1200
CGCCCAACAA	CGGGAGGCC	TGCCCCGAGC	TGGAAGAAGA	GGCTGAGTGC	GTCCCTGATA	1260
ACTGCGTCTA	AGACCAAGCG	CCCGCAGCCC	CTGGGGCCCC	CGGAGCCATG	GGGTGTGGGG	1320
GGCTCTGTGG	CAGGCTCATG	CTGCAGGCGG	CCGAGGCACA	GGGGGTTTGG	CGCTGCTCCT	1380
GACCGCGGTG	AGGCCGCGCC	GACCATCTCT	GCACTGAAGG	GCCCTCTGCT	GGCCGCGCAG	1440
GGCATGGGGA	AACAGGCTTC	TCTTTTCCCA	ACCTTGCTTC	TGAGGGGCCC	CGGTGTCCCG	1500
TCTGCTCTCA	CGCTCTCTCT	CCTGCAAGAT	AAAGTCATCC	CCAAGGCTCC	AGCTACTCTA	1560
AATATAGGTC	TCTTATAAAG	TTATTGCTGC	TCCAGGAGAT	TGTCCTTCAT	GTCCAGGGGG	1620
CCTGGCTCCC	ACGTGGTGGC	AGATACCTCA	GACCTGGTGC	TCTAGGCTGT	GCTGAGCCCA	1680
CTCTCCGAG	GGCGCATCCA	AGCGGGGGCC	ACTTGAGAAG	TGAATAAATG	GGCGGTTTTC	1740
GGAGCGGTCA	GTGTTTCCAT	GTTATGGATC	TCTCTGCGTT	TGAATAAAGA	CTATCTCTGT	1800
TGCTCAC						

Seq ID NO: 132 Protein sequence
Protein Accession #: NP_036577.1

70
75

1	11	21	31	41	51	
MENPSPAAL	AKALCALILA	TLGAACQPLG	GESICBARAP	AKYSITFTGK	WSQTAFPKQY	60
PLFRPFAQWS	SLIGAABSD	YSMRKNQYV	SNGLRDFAR	GEAWALKKEI	EAAGEALQSV	120
HAVPFAPAVP	SGTQTSAREL	EVQRHSLVS	FVVRIVPSFD	WFGVDSLDL	CGDRNREQA	180
ALDLYFYDAG	TDSGFTFSSP	NFATIPQDTV	TEITSSSPSH	PANSFFYDRL	KALPPIAKVT	240
LVRLRQSPRA	FIPPAFVLPS	RDNEIVDSAS	VPETFLDCFV	SLWSWGLCG	GHCGRIGTKS	300
RTRYVRVQPA	NNGSPCPBLE	EEAECPVDNC	V			

Seq ID NO: 133 DNA sequence
Nucleic Acid Accession #: NM_019894
Coding sequence: 1..1314

80

1	11	21	31	41	51	
ATGTTACAGG	ATCCTGACAG	TGATCAACCT	CTGAACAGCC	TGATGTCAA	ACCCCTGCGC	60

5
10
15
20

```

AAACCCCGTA TCCCATGGA GACCTTCAGA AAGGTGGGGA TCCCATCAT CATAGCACTA 120
CTGAGCCTGG CSAGTATCAT CATGTGGTGT GTCTTCATCA AGGTGATTCT GGATAAATAC 180
TACTTCTCTCT CGGGGCGAGCC TCTCCACTTC ATCCCGAGGA AGCAGCTGTG TGACGGAGAG 240
CTGGACTGTCT CCTTGGGGGA GGAAGAGGAG CACTGTGTCA AGAGCTTCCC CGAAGGGCCT 300
GCAGTGGCAG TCCBCTCTCT CAGGACCGA TCCACACTGC AGGTGCTGGA CTGGGCCACA 360
GGGAAGTGGT TCTCTGCTGT TTTGACAAAC TTCACAGAAG CTCTGCTGTA GACAGCCTGT 420
AGGCAGATGG GCTACAGCAG CAACCCCACT TTCAGAGCTG TGGAGATTGG CCCAGACCAG 480
GATCTGGATG TTGTGAAAT CACAGAAAAC AGCCAGGAGC TTCGATGCG GAACTCAAGT 540
GGGCCCTGTC TCTCAGGCTC CCTGGTCTCC CTGCACTGTC TTGCTGTGG GAAGAGCCTG 600
AAGACCCCCC GTGTGGTGGG TGGGAGGAG GCCTCTGTGG ATTCTTGGCC TTGGCAGGTC 660
AGCATCCAGT ACGACAAACA GCAGTCTGT GTAGGGAGCA TCCTGGACCC CCAGTGGGTC 720
CTACAGGAGG CCCACTGCTT CAGGAAACAT ACOGATGTGT TCAACTGGA GGTGCGGGCA 780
GGCTCAGACA AACTGGGCGA CTTCCTATCC CTGGCTGTGG CCAAGATCAT CATCATGAA 840
TTCAACCCCA TGTACCCCAA AGACAATGAC ATCGCCCTCA TGAAGCTGCA GTTCCCACTC 900
ACTTCTCAGC GCACAGTCAG GCCATCTGT CTGCCCTTCT TTGATGAGGA GCTCACTCCA 960
GCCACCCACC TCTGGATCAT TGGATGGGGC TTTACGAAGC AGAATGGAGG GAAGATGTCT 1020
GACATACCTG TCGAGGGGTC AGTCCAGGTC ATTGACAGCA CACGGTGCAA TGACAGCAT 1080
GGTACCCAGG GGGAAAGTAC CAGAAAGATG ATGTGTGACG GCATCCCGGA AGGGGGTGTG 1140
GACACCTCGC AGGGTGAAC TGGTGGGCC CTGATGTACC AATCTGACCA GTGGCATGTG 1200
GTGGGCATCG TTAGCTGGGG CTATGGCTGC GGGGGCCCGA GCACCCAGG AGTATACACC 1260
AAGGTCTCAG CCTATCTCAA CTGATCTAC AATGTCTGGA AGGCTGAGCT GTAA

```

Seq ID NO: 134 Protein sequence

Protein Accession #: NP_063947.1

25
30
35

```

1 11 21 31 41 51
| | | | |
MLQPDSDQF LNSLDVFLR KPRIPMETFR KVGIPILIAL LSLASIIIV VLIKVILDKY 60
YFLCGQFLRF IPRKQLCDGE LDCPLGEDEE HCVKSPFEGP AVAVRLSKDR STLQVILDSAT 120
GNWFACTPN PTEALAEFAC RQMYSKSKPT FRAVEIGPDQ DLDVVEITEN SQELRMENSS 180
GPCLSGSLVS LHCLACGKSL KTRPVVGEE ASVDSWPNQV SIQYDKQHVC GGSILDPHNV 240
LTAABCFRKH TDVFNWVRRA GSKLGSFPF LAVAKIILF FNPMPKDNQ IALMKLQFPL 300
TFSGTVPFIC LPFFDEBLTP AITPLIIGWG FTKQNGGKMS DILLQASVQV IDSTRCNADD 360
AYQGEVTERK MCAGIFPESGV DTCQEDSGGP LMYQSDQWHV VGIWSWGYGC GGPSTPGVYT 420

```

Seq ID NO: 135 DNA sequence

Nucleic Acid Accession #: NM_003045

Coding sequence: 148..2037

40
45
50
55
60
65
70
75
80

```

1 11 21 31 41 51
| | | | |
CGATCTGCC GAGGCCCGC GCGGCCCGC TTGGATTCTG AAACCTTCCT TGTATCCCTC 60
CTGAGACATC TTGTCTGCAA GATCGAGGCT GTCTCTGGT GAGAAGGTGG TGAGGCTTCC 120
CGTCATATTC CAGCTCTGAA CAGCAACATG GGGTGCAAAG TCCTGCTCAA CATTTGGCAG 180
CAGATGCTGC GGGGGAAGGT GGTGACTGT AGCCGGGAGG AGACGCGGCT GTCTGCTGAC 240
CTGAACACTC TTGATCTGGT GGCCTCTGGG GTGGGCAGCA CACTGGGTGC TGGTGTCTAC 300
GTCTCTGGCT GAGCTGTGGC CGTGAGAAAT GCAGGCCCTG CCATTGTCTAT CTCCTTCTGC 360
ATGCTGTGGC TGGCTCTAGT GCTGGCTGGC CTGTGCTATG GCGAGTTTGG TGTCTGGGTC 420
CCCAAGAGCG GCTCAGCTTA CCTCTACAGC TATGTACCG TTGGAGAGCT CTGGCCCTTC 480
ATCACCGGCT GGAACCTTAAT CCTCTCTAC ATCATCGTA CTTCAAGGT AGCGAGGGCC 540
TGGAGCGGCA CCTTGAGGGA GCTGATAGGC AGACCATCG GGGAGTTCTC ACGGACACAC 600
ATGACTCTGA ACGCCCGCGG CGTCTGGCT GAAACCCCG ACATATTGCG AGTGATCATA 660
ATTCTCATCT TGACAGGACT TTAACTCTT GGTGTGAAG AGTCGGCCAT GGTCAACAAA 720
ATATTCACTT GTATTAAAGT CCTGCTCTG GGCTTCATAA TGGTGTGAGG ATTGTGAAA 780
GGATCGTTA AAAACTGGCA GCTCACGGAG GAGGATTTG GGAACACATC AGGCCGTCTC 840
TGTATTGAACA ATGACACAAA AGAAGGGAAG CCCGGTGTG GTGGATTGAT GGCCTTGGG 900
TTCTCTGGTG TCTCTGGG GGCAGCGACT TGCTTCTATG CCTTCGTGGG CTTTGACTGC 960
ATGCCACCA CAGGTGAAGA GGTGAAGAAC CCACAGAAGG CCATCCCGGT GGGGATGTG 1020
GGTCTCTCT TGTATGCTT CATGCGCTAC TTGGGGGTG CAGCTGCCCT CACGCTCATG 1080
ATGCCCTACT TCTGCTGGA CAATAACAGC CCCTGCCC GCGCTTTAA GCAGTGGGC 1140
TGGGAAGGTG CCAAGTACGC AGTGGCCGTG GGTCTCTCT GGCCTCTTC CCGCAGTCTT 1200
CTAGGTTCOA TGTTTCCCAT GCCTCGGTT ATCTATGCCA TGGCTGAGGA TGGACTGCTA 1260
TTTAAATCT TAGCCAAGT CAATGATAGG ACCAAACAC CAATAATGCG CACATTAGCC 1320
TGGGTGCGG TTGCTGCTG GATGGCCTTC CTCTTGACC TGAAGGACTT GGTGGACCTC 1380
ATGTCCATTG GCATCTCTCT GGCCTACTCG TTGGTGGCTG CCTGTGTGTT GGTCTTACGG 1440
TACCAGCCAG AGCAGCTTAA CCTGGTATAC CAGATGGCCA GTACTTCCGA CGAGTTAGAT 1500
CCAGCAGACC AAAATGAATT GGCAGCACC AATGATCCC AGCTGGGGT TTTACAGAG 1560
GCAGAGATGT TCTCTTGAA AACCATACTC TCACCCAAA ACATGGAGCC TTCCAAAATC 1620
TCTGGGCTAA TTGTGAACAT TTCAACAGC CTATAGCTG TTCTCATCAT CACCTCTGCG 1680
ATTGTGACCB TGCTTGGAGG GAGGGCTCTC ACCAAAGGGG CGCTGTGGGC AGTCTTCTG 1740
CTCGACGGGT CTGCCCTCT CTGTGCTGTG GTCAAGGGCG TCATCTGGAG GCACCCGAG 1800
AGCAAGACCA AGCTCTCAT TAAAGTTCCC TTCTGCCCAG TGCTCCCAT CCTGAGCATC 1860
TTCTGTGAAG TGGCTCTCAT GATGACCTG GACCAAGGCA CCTGGGTCCG GTTGTCTGTG 1920
TGGATGCTGA TAGGCTCTAT CATCTACTT GGCATGCCC TGTGGCAGC CGAGGAGGCG 1980
TCCCTGGATG CCGACCAAGC AAGGACTCT GACGGCACT TGGACCATG CAAGTGACGC 2040
ACAGCCCGC CCGCGAGG TGGCAGCAG CCGAGGGAC GCCCCAGAG GACCGGGAGG 2100
CACCCACCC TCCACACAG TGCACAGAA ACCACCTGCG TCACACCCCT CACTGCA

```

Seq ID NO: 136 Protein sequence

Protein Accession #: NM_003045

1 11 21 31 41 51
| | | | |

5 MGCKVLLNIG QQMLRRKVVD CSREETRLSR CLNTFDLVAL GVGSTLGAGV YVLGAVARE 60
 NAGPAIVISF LIAALASVLA GLCYGSEFGAR VPKTGSAYLY SYVTVBELWA FITGNLILS 120
 YIIGTSSVAR AWSATFDELI GRPIGEPSRT HMTLNAPGVL AENPDIPAVI IILILTGLLT 180
 LGVKESAMVN KIFTCLNVLV LGFIMVSGFV KGSVKWQLT EEDFGNTSGR LCLNMDTKEG 240
 KPGVGGFMFP GFSGVLSGAA TCFYAFVGF D CIATTGEEVK BPQKAIPVGI VASLLICFIA 300
 YFGVSAALTL MMEPYCLDNN SPLPDAFKHV GWEGAKYAVA VGSCLALSAS LLSGMPFMPR 360
 VIYAMAEDGL LFKPLANVND RTKTPITATL ASGAVAAMVA FLFDLKLVD LMSIGTLAY 420
 SLVAACVLVL RYQPEQPNLV YQMASTSEDL DPADQNELAS TNDSQLGLFP EAMFSLKTI 480
 LSPKNMEPEK ISGLIVNIST SLIAVLIITF CIVTVLGREA LTKGALWAVF LLASGALLCA 540
 10 VVTGVIWRQP ESKTKLSFKV PFLFVLPTLS IFNVVILMMQ LDQGTWVRFA VWMILGFIIT 600
 FGYGLMHSSE ASLDADQART PDGNLDQCK

Seq ID NO: 137 DNA sequence

Nucleic Acid Accession #: NM_032044.1

Coding sequence: 182..658

15
 1 11 21 31 41 51
 20 AAGATATAAA AGCTCCAGAA ACGTTGACTG GGACCACTGG AGACACTGAA GAAGGCAGGG 60
 GCCCTTAGAG TCTTGGTTGC CAAACAGATT TGCAATCAAA GGAGAACCCA GGAGTTTCAA 120
 AGAAGCGCTA GTAAGGTCTC TGAGATCCTT GCACTAGCTA CATCTCAGG GTAGGAGGAA 180
 GATGGCTTCC AGAAGCATGC GGCCTGCTCT ATTGCTGAGC TGCCCTGGCCA AAACAGGAGT 240
 CCGGGTGTAT ATCATCATGA GACCCAGCTG TGCTCTGGA TGGTTTACC ACAAGTCCAA 300
 25 TTGCTATGGT TACTTCAGGA AGCTGAGGAA CTGGTCTGAT GCGGAGCTCG AGTGTCASTC 360
 TTACCGAAAC GGAGCCCACT TGGCATCTAT CTTGACTTTA AAGGAAGCCA GCACCATAGC 420
 AGATACATA AGTGGCTATC AGAGAAGCCA GCGATATGG ATTGGCTGAC ACGACCCACA 480
 GAAGAGGCAG CAGTGGCAGT GGAATGTATG GGCATGTAT CTGTACAGAT CTTGGTCTGG 540
 CAGTCCATG GGTGGGAACA AGCACTGTGC TGAGATGAGC TCCAATAACA ACTTTTAAAC 600
 30 TTGAGCAGC AACGAATGCA ACAGCGCCA ACACCTCTCT TGCABSTACC GACCATAGAG 660
 CAGAATCAA GATCTGCTA ACTCCTGCAC AGCCCGCTCC TCTTCTTTC TGCTAGCCTG 720
 GCTAATCTG CTCATTATT CAGAGGGGAA ACCTAGCAAA CTAGAGTGA TAAGGGCCCT 780
 ACTACACTGG CTTTTTAGG CTTAGAGACA GAAACTTTAG CATTGGCCA GTAGTGGCTT 840
 CTAGCTCTAA ATGTTTGCCG CGCCATCCCT TTCCACAGTA TCCCTCTTCC CTCTCCCTCT 900
 35 GTCTCTGGCT GTCTCGAGCA GTCTAGAAGA GTGCTATCC AGCCTATGAA ACAGCTGGGT 960
 CTTTGGCCAT AAGAAGTAAA GATTGAAGA CAGAAGGAAG AAACCTCAGG GTAGGCTTCT 1020
 AGACCCCTTC AGCTTCTACA CCCTTCGACC CTCTCTCCAT TGCTTGCACC CCACCCGAGC 1080
 CACTCAACTC CTGCTTGTTT TTCTTTGGC CATAGGAAGG TTTACCACTA GAATCCTTGC 1140
 TAGGTGTATG TGGGCCATAC ATTCCCTTAA TAAACCATG TGTACATAAG AAAAAAANA

Seq ID NO: 138 Protein sequence

Protein Accession #: NP_114433.1

45 1 11 21 31 41 51
 MASRSMRLLL LLSCLAKTGV LGDIIMRPSK APGWFYHKSN CYGYFRKLEN WSDAELEBQS 60
 YGNAGRLASI LSLKEASTIA EYISGYQRSQ PIWIGLEDPQ KRQWQWIDG AMYLYRSWSG 120
 KSMGKNKHA EMSNNNFLT WSSNECNKRQ HFLCKYRP

Seq ID NO: 139 DNA sequence

Nucleic Acid Accession #: XM_051860.2

Coding sequence: 52..3042

55 1 11 21 31 41 51
 GCTCACCAG GAAAAATATG CAATCGTCCC ATTGATATAC AGGCCACTAC AATGGATGGA 60
 GTTAACCTCA GCACCGAGGT TGTCTACAAA AAAGGCCAGG ATTATAGGTT TGCTTGCTAC 120
 GACCGGGGCA GAGCCTGCGG GAGCTACCGT GTACGGTTCC TCTGTGGGAA GCCTGTGAGG 180
 CCAACACTCA CAGTCAACAT TGACACCAAT GTGAACAGCA CCACTCTGAA CTGTGGAGAT 240
 AATGTACAGT CATGGAACC TGGAGATACC CTGGTCATTC CCAGTACTGA TTACTCCATG 300
 60 TACCAGGAG AAGAGTTCCA GTTGCTTCCC TGCAATCCT GCGCCGCCAA CCAAGTCAAA 360
 GTGGCAGGGA AACCAATGTA CCTGCACATC GGGGAGGAGA TAGACGGCGT GGACATGCGG 420
 GCGGAGGTTG GGCCTCTGAG CCGGAACATC ATAGTATGAG GGGAGATGGA GGACAAATGC 480
 TACCCCTACA GAACACCATC CTGCAATTPC TTTGACTTCG ATACCTTTGG GGGCCACATC 540
 AAGTTTGCTC TGGGATTTAA GGCAGCACAC TTGGAGGGCA CGGAGCTGAA GCATATGGGA 600
 65 CAGCAGCTGG TGGGTCACTA CCGCATTCAC TTCCACCTGG CCGGTGATGT AGACGAAAGG 660
 GGAGGTATG ACCACCCAC ATACATCAGG GACCTCTCCA TCCATCATAC ATTCTCTGCG 720
 TGCGTCACAG TCCATGGCTC CAATGGCTTG TTGATCAAGG ACGTTGTGGG CTATAACTCT 780
 TTGGGCACTT GCTTCTTCAC GGAAGATGGG CCGGAGGAAC GCAACACTTT TGACCACTGT 840
 70 CTTGGCCTCC TTGTCAAGTC TGGAAACCTC CTCCCTCGG ACCGTGACAG CAAGATGTGC 900
 AAGATGATCA CAGGAGACTC CTACCCAGG TACATCCCA AGCCCAAGCA AGACTGCAAT 960
 GCTGTGTTCA CTTCTGGAT GGCAATCCC AACACAACC TCATCAACTG TGCCGCTGCA 1020
 GGATCTGAGG AAACCTGATT TTGGTTTAT TTGACCCAG TACCAACGGG CCCCTCCGTG 1080
 GGAATGACT CCCAGGTTA TTAGAGCAC ATTCCACTGG GAAATTTCTA TAACAACCGA 1140
 75 GCACATTCCA ACTACCGGGG TGGCATGATC ATAGACAACG GAGTCAAAAC CACCGAGGCC 1200
 TCTGCCAAGG ACAAGCGGCC GTCTCTCTCA ATCATCTCTG CCAGATACAG CCCTCACCAG 1260
 GAGCGGAGC CGTGAAGCC CCGGAGGCCG GGCATCATCA GACACTTCAT TGCCATCAAG 1320
 AACCAGGACC ACGGGGCGTG CTTGGCGGCG GGGGATGTGT GCTTGGACAG CTGCGGTTT 1380
 GCTGACATG GCATTGGCCT GACCTTGGCC AGTGGTGGAA CTTTCCCGTA TGACACAGGC 1440
 80 TCCAGCAGG AGATAAGAA CAGCTTGTIT GTTGGCGAGA GTGGCAACGT GGGGACGGAA 1500
 ATGATGGACA ATAGGATCTG GGGCCCTGGC GGCCTGGACC ATAGCGGAAG GACCCTCCCT 1560
 ATAGGCCAGA ATTTCCAT TAGAGGAATT CAGTTATATG ATGGCCCAT CAACATCCAA 1620
 AACTGCACIT TCGAAGASTT TGTGGCCCTG GAGGGCCGGC ACACCAAGCG CTTGGCCTTC 1680
 CGCCGGAATA ATGCTTGGCA GAGCTGCCCC CATAACAACG TGACCGGCAT TGCTTTGAG 1740
 GACGTTCCGA TTACTTCCAG AGTGTCTCTC GAGAGGCTG GGCCTTGGTT CAACAGCTG 1800

5
10
15
20
25
30
35
40
45
50
55
60
65

GACATGGATG GGGATAAGAC ATCTGTGTTC CATGACGTGG ACBGGCTCCGT GTCGAGTAC 1860
 CCTGGCTCCT ACCTCAGGAA GAATGACAACT TGGCTGGTCC GGCACCCAGA CTGCATCAAT 1920
 GTTCCCGACT GGAGAGGGGC CATTTGCACT GGGTGTCTATG CACAGATGTA CATTCAAGCC 1980
 TACAGACCA GTAACCTGGG AATGAAGATC ATCAAGAATG ACTTCCCCAG CCACCCCTCTT 2040
 TACCTGGAGG GGGCGCTCAC CAGGAGCACC CATTACCAGC AATACCAACC GGTGTCTACC 2100
 CTGCAGAAAG GCTACACCAT CCATCTGGAC CAGACGGCCC CCGCCGAACT CGCCATCTGG 2160
 CTCAATCACT TCAACAAGGG CGACTGGATC CGAGTGGGGC TCTGCTACCC GCGAGGCACC 2220
 ACATTCTCCA TCTCTCGGA TGTTCACAAAT CGCTGTCTGA AGCAAACTGC CAAGACGGGC 2280
 GTCTTCTGTA GACCTTTGCA GATGACAAA GTGAGCAGA GCTACCCTGG CAGGAGCCAC 2340
 TACTACTGGG AGGAGACTC AGGGCTGTTG TTCTGAAAGC TGAAGCTCA GAACGAGAGA 2400
 GAGAAAGTTG CTTTCTGTCT CATGAAAGGC TGTGAGAGGA TAAAGATTAA AGCTCTGATT 2460
 CCAAGAAAGC CAGGCGTCAG TGACTGCACA GCCACAGCTT ACCCAAGTT CACCGAGAGG 2520
 GCTGTCTGAT AGTGTCCGAT GCCCAAGAG CTCTTTGGTT CTCAGCTGAA AACAAAGGAC 2580
 CATTTCTTGG AGGTGAAGAT GAGGAGTTCC AAGCAGCACT TCTTCCACCT CTGGAACGAC 2640
 TTCTCTTACA TTGAAGTGGG TGGGAAGAGG TACCCAGTT CGGAGGATGG CATCCAGGTG 2700
 GTGTGTGATT ACGGGAACCA AGGGCGCGTG GTGAGCCACA CGAGCTTCAG GAACTCCATT 2760
 CTGCAAGGCA TACCATGGCA GCTTTTCAAC TATGTGGCGA CCATCCCTGA CAATTCCATA 2820
 GTGCTTATGG CATCAAGGGG AAGATACGTC TCCAGAGGCC CATGGACCAG AGTGTGGAA 2880
 AAGCTTGGGG CAGACAGGGG TCTCAAGTTG AAAGAGCAAA TGGCATTCTT TGGCTTCAA 2940
 GGCAGCTTCC GCGCCATCTG GGTGACACTG GACACTGAGG ATCACAAGC CAAATCTTTC 3000
 CAAGTTGTGC CCATCCCTGT GGTGAAGAG AAGAAGTTGT GAGGACAGCT GCCGCCCGGT 3060
 GCCACCTCGT TACGACTGAT GACGGTGAAT CTTGGCAGCA GACCACTGGG GGTATGGCTGG 3120
 GTCCCTCAGC CCTGTCCAGC AGCTGCTGGG GAAGGCCGTG TTTACGCCCT GATGGGCCAA 3180
 GGGAAAGGCTA TCAGAGACCC TGGTGTCTGC ACCTGCCCTT ACTCAAGTGT CTACCTGGAG 3240
 CCGCTGGGGC GGTGCTGGCC AATGCTGGA ACATTCACCT TCTGTGAGCC TCTTGGGTGC 3300
 TTCTCTCTTA TCTGTGCTTC TTGAGTGGG GTTGGGGAC CATATCAGGA GACCTGGGT 3360
 GTGCTGACAG CAGGCTGACA CTTTGGCAGG AGCCCTGACC CAGCTAGGAG GTAGTCTGGA 3420
 GGGCTGTGTA TTCACAGATC CCAATGGTCT TCAGCAGACA AGTGAGGCTG GTAAATGTAG 3480
 GAGAAAGAGC TTGGGCTTGA AGGAAATCTT TACTCCTGTA AGCAAGAGCC AACCTCACAG 3540
 GATTAGAGGC TGGGCTAGAA CTGGCTATCC TTGGGGAAGA GGCAGGCCCT GCCTCTGGCC 3600
 GTGTCCACCT TTGAGGAGAC TTGAGTGGC AGGTTTGGAC TTGGACTAGA TGACTCTCAA 3660
 AGGCCCTTTT AGTCTCTGAA TTCCAGAAAT CTGCTGCAAT TCACATGGA CCTGGAACCC 3720
 AACAGTTCAAT GATATTCAC TGATATCCAT GATGCTGGGT GCCCAGCCGC ACACGGGATG 3780
 GAGAGGTGAG ACTAATATCC TAGCTTGAGG GGTCTGCACT CCAGTAGGAC AGGCACTCAG 3840
 GTTCACTGTC ACTGCAATGC CAGGTGGAGA AATCACAGAG AGGTAAATAG GAGGCCAGTG 3900
 CCATTTGAGA GGGGAGGCTC AGGAAGGCTT CTGCTTACA GGAATGAAGG CTGGGGGCAT 3960
 TTTGCTGGGG GAGAGTGGAG CAGCCTCTGG AATGGCTCAG GGATTGAGC CTCCCTGGCG 4020
 CTGCTCTGCT AAGCTGTGTA CTACGGGGTC GCGCTTTGCT CAGCTCTCTC TGGCCCACTC 4080
 ATGATGAGGA AGTGTGCTCA GAGGGGAGCA ATGGGCTTTG CTGCTTATGA GCACAGAGGA 4140
 ATTCACTCCC CAGGCGAGCC TGCTCTGAC TCCAGAGGGG TGAAGTCCAC AGAAGTGAGC 4200
 TCTGCTCTTA GGGCTCTCAT TGCTCTTCTT CCAGGGAATC GAGCAGAGGG GGCTCCAGG 4260
 AGACCTGAGA TGTGCTGCTA CTCCCTGCGC CTGGGATTTT AGAGCTGGAA ATATAGAAAT 4320
 TATCTAGCCC AAGCCTTCA TTTTAACAGA TGGGGAAGT GAGCCGCCAA GATGGGAAAG 4380
 AACCAACAG CTAAGGGAGG GCGTGGGGAG CCCCACCTTA GCGCTTGTG CCACACCACA 4440
 TTGCTCTAAC AACCGGCCCC AGAGTGCCCA GGCALTCTTG AGGTAGCTTC TGGAAATGGG 4500
 GACAGTCCC CTCGAGAGGA AGGAAATGAC TAGAGTAGAA TGACAGCTAG CAGATCTCTT 4560
 CCGCTCTGCT CCGAGCGCAT ACAGACCCGC CCTCCCTTGG GTGTGGGCGG TCCCTGTGGC 4620
 CTCTACTTTG TTCACTACCT GTGAGCCAG CCGTGGGTGA CAGTAGCTGC AACTCCCCAT 4680
 TGGTGTACG TGCTCTCTCT GTCTCTGAG CTCTACAGGT GAGGCCGAGC AGAGGGAGTA 4740
 GGGCTGCGCA TGGTTCTGGT GAGCCAAATT GCGTATCTT GGGTGTCTGA ACAGCTATTG 4800
 GGTCCACCCC AGTCCCTTTC AGCTGCTGCT TAATGCCCTG CTCTCTCCCT GGGCCACCTT 4860
 ATAGAGAGCC CAAAGAGGCT CTGTAAGAGG GAGAACTCTA TCTGTGTTT ATAATCTTGC 4920
 ACGAGGCACC AGAGTCTCCC TGGGTCTTGT GATGAACATC ATTTATCCCC TTCTCTGCC 4980
 CAACCAAAA CTCTTCTCTT CAAAGAGGGC CTGCTGGCTT CCTTCCACC CCTGTCACCC 5040
 ATGAGACTCG GTCCAGAGGT CCATTCGCCA GGTGGGAGCC AACTGTGACG GAGGTCTTTC 5100
 CCACCAACAA TCTTTGAGCT GCTGGGAGGT GACCATAGGG CTCTGCTTTT AAGATATAGG 5160
 CTGCTCTAAA GCGCAGAGCT ACAGGAAGGA CTCTTCCAG GGAGATTAGT GGTGATGGAG 5220
 AGGAGAGTTA AATGAGCTC ATGTCTCTCT TGTCCAGGT TTTGTTAGT TTTCACTCTT 5280
 CTAATGCAAG GGTCTCACAC TGTGAACAC TTAGGATGTG ATCACTTTCA GGTGGCCAGG 5340
 AATGTTGAAT GTCTTTGGCT CAGTTTATT AAAAAAGATA TCTATTGAA AGTTCCTAGA 5400
 GTGTACATA TGTTCACAG TACAGGATCT GTACATAAAA GTTCTTTCC TAAACCATTC 5460
 AOCAGAGGCC AATATCTAGG CATTTTCTTG GTAGCACAAA TTTCTTATT GCTTAGAAAA 5520
 TTGTCTCTCT TGTATTCTG GTTTGTAAGA CTTAAGTGAG TTAGCTCTT AAGGAAGCA 5580
 ACGCTCTCTT GAATGCTTG TCTTTTCTT GTTGGGAAA TAGCTGGTCC TTTTCTGGGA 5640
 GTTAGAGTA TAGAGTGTG GTATGTAAAC ATTTCTTGA GGCATCACA TGAACAAAGA 5700
 TATATTTCT ATTTATTAT TATATGTGCA CTTCAAGAAG TCACCTGTAG AGAAATAAG 5760
 AATTGTCTTA AATGTCAAAA AAAAAA AAAA AAAA

70 Seq ID NO: 140 Protein sequence
 Protein Accession #: XP_051860.2

75
80

1 11 21 31 41 51
 | | | | |
 MOGVNLSTEV VYKKGQDYRF ACYDRGRACR SYRVRFLCGK FVRPKLTVTI DTNVNSTILN 60
 LEDNVQSWKP GDTLVIASTD YSMYQAEERF VLPFRSCAPN QVKGAGKPMY LHIGEEIDGV 120
 DMRABVGLLS RNIIVMGEME DKCYFYRNHI CNFFDFDTFG GHIKFAFGFK AAHLEGTELK 180
 HMQQLVGVQY PIFHFLAGDV DERGGYDPPT YIRDLSEIHT FSRCTVHGS NGLLKVGVG 240
 YNSLGHCFPT EDGPSEKNTF DHCLGLNLXS GTLLPSDRDS KMKMTIGDS YRGYIPKPRQ 300
 DCNAVSTFWM ANPNNNLINC AAGSSESTGF WFIHNVPTG PSVGMYSPTY SEHIPLGKPY 360
 NNRAHNSYRA GNIIDNGVKT TEASAKDKRP FLRIISARYS PHQDADPLKP REPATIRHFI 420
 AYXNQDEGAM LRGGDVWLDG CRFADNGIGL TLASGGTFPY DGSKEIEN SLFVGESGNV 480
 GTEMDMNRIN GFGGLDEBGR TLPIGNPFI RGIQLYDGP I NQNCIFRKF VALEGRHTSA 540
 LAFLRLNNAQ SCPHNNVTGI APEDVPITSR VFGEKGPWF NQLDMGDKT SVFEDVDGSV 600

SEYPGSYLTK NDNLVLRHPD CINVPDWRGA ICSGCTAOMY IQAYKTSNLR MKIIKNDPPS 660
 HPLYLEGALT RSTHYQQYQP VVTLQKGYTI HWDQTAPAEI AIWLINPNKG DWIRVGLCYP 720
 RGTFFSILSD VHNRLKLQTS KTGVPVRLIQ MDKVEQSYFG RSHYYNDEDS GLLFLKLKLAQ 780
 NEREKFAFCS MKGCEKIKK ALIPKNAGVS DCTATAYPKF TERAVVDVPM PKLFLGSQLK 840
 TKDHFLVFKM ESSKQHFHFL WNDPAYIEVD GKYPSSSEDG IQVVVIDGNQ GRVVSSTSRF 900
 NSILQGIPWQ LFNYYVATIPD NSIVLMASKG RYVSRGPWTR VLEKLGADRG LKLKEQMAFV 960
 GFKSGFRPIW VTLDTEDHKA KIPQVVPFV VKKKKL

Seq ID NO: 141 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 261..2861

1 11 21 31 41 51
 15 GAGCTAGCGC TCAAGCAGAG CCCAGCGCGG TGCTATCGGA CAGAGCGCTGG CSAGCGCAAG 60
 CGGCBGCGGG AGCCAGCGGG GCTGAGCGCG GCCAGGGTCT GAACCCAGAT TTCCCGAGCT 120
 AGCTACCACT CGGCTTGGCC AGCCCCCGGG AGCTCGCGGC GCCTGGCGGT CAGCGACCAG 180
 ACGTCCGCGG CCGCTGCGCT CCTGGCCCGC GAGGCGGTAC ACTGTCTCGG CTACAGACCC 240
 20 AGAGGGAGCA CACTGCCAGG ATGGGAGCTG CTGGGAGGCA GGACTTCCCT TTCAAGGCCA 300
 TGCTGACCAT CAGCTGGCTC ACTCTGACCT GCTTCCCTGG GGCCACATCC ACAGTGGCTG 360
 CTGGGTGCCC TGACCCAGAG CCTGAGTTGC AACCTTGGAA CCTTGGCCAT GACCAAGACC 420
 ACCATGTGCA TATCGGCCAG GGCAGAGAC TGCCTGCTCAC CTCTCTGCCC ACGGTCTATT 480
 CCATCCACAT CTCAGAGGGA GGCAAGCTGG TCATTAAAGA CCACGACGAG CCGATTGTCT 540
 25 TGGGAACCCG GCACATCCCT ATTGACAACG GAGGAGAGCT GCATGCTGGG AGTGCCTCTG 600
 GCGCTTTCCA GCGCAATTTT ACCATCATTT TGTATGGAAG GCGTGTATGA GGTATTCAAG 660
 CGGATCCCTA CTATGGTCTG AAGTACATIG GGGTGTGTA AGGAGCGCT CTGTAGTTGC 720
 ATGACAGAA AAAGCTCTCC TGGACATTT TGAACAAGAC CCTTCACCCA GGTGGCATGG 780
 CAGAAGGAGG CTATTTTTTT GAAAGGAGCT GGGGCCACCG TGGAGTTATT GTTCATGTCA 840
 30 TCGACCCCAA ATACGCCACA GTCATCCATT CTGACCGGTT TGACACCTAT AGATCCAGA 900
 AAGAGAGTGA ACGTCTGGTC CAGTATTTGA ACGCGGTGCC CGATGGCAGG ATCCTTCTCG 960
 TTGCAGTGAA TGATGAAGGT TCTCGAAATC TGGATGACAT GGCCAGGAAG GCGATGACCA 1020
 AATTTGGGAG CAAACACTTC CTGACCTTGG GATTTAGACA CCTTTGGAGT TTTCTAATCG 1080
 TGAAGGAAAT TCCATCATCT TCAGTGGGAG ACCATATTGA ATATCATGGA CATCGAGGCT 1140
 35 CTGCTGCTGC CGGGTATTC AAATTTGTTCC AGACAGAGCA TGGCGAATAT TTCAATGTTT 1200
 CTTTGTCCAG TGAATGGGTT CAAGACGTGG AGTGGACGGA GTGGTTCGAT CATGATAAAG 1260
 TATCTCAGAC TAAAGGTGGG GAGAAATTT CAGACCTCTG GAAAGCTCAC CCAGGAAAAA 1320
 TATGCAATCG TCCCATTTGAT ATACAGGCCA CTACAATGGA TGGAGTTAAC CTCAGCACCG 1380
 AGGTTGCTCA CAAAAAGGC CAGGATTATA GGTTTGCTTG CTACGACCGG GGCAGAGCCT 1440
 40 GCGGAGGCTA CCGTGTACGG TTCTCTGTG GGAAGCCTGT GAGGCCCAA CTCACAGTCA 1500
 CCATTGACAC CATATGAAAC AGCACCATTC TGAACCTTGA GGATAATGTA CAGTCATGGA 1560
 AACCTGGAGA TACCTTGGTC ATTGCCAGTA CTGATTACTC CATGTACAGG GCAGAGAGAT 1620
 TCCAGGTGCT TCCCTGCGAG TCCCTGCGCC CCAACCCAGT CAAAGTGGCA GGGAAACCAA 1680
 45 TGAGCCGAAA CATCTGGGAG GAGATAGAGG GCGTGGACAT GCGGGCGGAG GTTGGGCTTC 1740
 ACATCTGCAA TTCTTTGAC TTGCATACCT TTGGGGGCGA CATCAAGTTT GCTCTGGGAT 1800
 TTAAGGCAGC ACACCTGGAG GGCACGGAGC TGAAGCATAT GGGACAGCAG CTGGTGGGTC 1860
 AGTACCGGAT TCACTTCCAC CTGGCCGGTG ATGTAGACGA AAGGGGAGGT TATGACCCAC 1920
 CCACATACAT CAGGGACCTC TCCATCCATC ATACATTTCT TCGCTGCTC ACAGTCCATG 2040
 50 GCTCCATGCG CTGTGTGATC AAGGACGTTG TGGGCTATAA CTCTTTGGGC CACTGCTTCT 2100
 TCACGGAAGA TGGGCGGAG GAAACGACAA CTCTTGACCA CTGCTTGGGC CTCTTGTCTA 2160
 AGTCTGGAAC CCTCTTCCCC TCGGACCTGT ACAGCAAGAT GGCAGAGATG ATCAGAGAGG 2220
 ACTCTTACCC AGGGTACATC CCCAAGCCCA GSCAAGACTG CAATGCTGTG TCCACCTTCT 2280
 GGATGGCCAA TCCCAACAA ACCTCATCA ACTGTGCGCG TGAGGATCT GAGGAAACTG 2340
 55 GATTTTGGTT TATTTTTCAC CACGTACCAA CCGGCCCTTC CBTGGGAATG TACTCCCCAG 2400
 GTTATTGAGA GCACATTCCA CTGGGAAAT TCTATAACAA CCGAGCACAT TCCAACTACC 2460
 GGGCTGGCAT GATCATAGAC AAACGAGTCA AAACCAAGA GGCCTTGCC AAGGACAGC 2520
 GGCCTTCCCT CTCAATCATC TCTGCCAGAT ACAGCCCTCA CCAGGACGCC GACCGCTGA 2580
 AGCCCGGGGA GCGCGCCATC ATCAGACACT TCATTGCTTA CAAGAACCAG GACCACCGGG 2640
 60 CCTGGCTGCG CCGCGGGAT GTGTGGCTGG ACAGCTGCCA TTTAGAGGG GAGGCTCAGG 2700
 AAGGCTTCTT GCTTACAGGA ATGAAGGCTG GGGGCATTTT GCTGGGGGGA GATGAGGCAG 2760
 CCTCTGGAAT GGCCTAGGGA TTCAGCCCTC CCTGCGGCTG CCGCTGGAAG CTGGTGAATA 2820
 CGGCTGCGCC CTTTGTCTAC GTCTCTCTGG CCGACTCATG ATGGAGAAGT GTGGTCAGAG 2880
 GGGAGCAATG GCGTTTCTG CTTATGAGCA CAGAGGAATT CAGTCCCGAG GCAGCCCTGC 2940
 65 CTCTGACTCC AAGAGGGTGA AGTCCACAGA AGTGAGCTCC TGCCTTAGGG CCTCATTTGC 3000
 TCTTCATCCA GGGAACTGAG CACAGGGGCG CTCCAGGAGA CCTAGATGT GCTGTACTC 3060
 CTTGCGCCTG GGAATTTCAG GCTGGAAATA TAGAAATAT CTAGCCCAA GCTTCTATT 3120
 TAAAGATGAG GGAAGTGAAG CCCCAAGAT GGGAAAGAAC CACACAGCTA AGGGAGGGCC 3180
 TGGGAGGCC CACCTAGCTT CTTGCTGCCA CACCACATTG CCTCAACAC CCGCCCCAGA 3240
 70 GTGCCCAGGC ACTCCTGAGG TAGCTTCTGG AAATGGGGAC AAGTCCCTTC GAGGGAAGG 3300
 AAATGACTAG AGTAGAATGA CAGCTAGCAG ATCTCTTCCC TCTGCTCCC AGCCACACA 3360
 AACCCGCCCT CCGCTTGGTG TTGGCGGTCC CTGTGGCTTC CACTTGTTC ACTACCTGTC 3420
 AGCCAGCCCT GGGTGCACAG TAGCTGCRAC TCCCATTTGG TGCTACCTGG CTCTGCTGTC 3480
 TCTGCAGCTC TACAGGTGAG GCCCAGCAGA GGGAGTAGGG CTGCCCATGT TTTGTTGAG 3540
 75 CCAATTTGGC TGATCTTGGG TGTCTGAACA GCTATTGGGT CCACCCAGT CCGTTTCAGC 3600
 TGCTGCTTAA TGCCCTGCTC TCTCCCTGGC CCACCTTATA GAGAGCCCAA AGAGCTCCTG 3660
 TAAAGGGGAG AACTCTATCT GTGGTTTATA ATCTTGACAG AGGCACAGA GTCTGCCCTG 3720
 GTCTGTGAT GAACACTATT TATCCCTTTT CTTGCCCAA CCACAACTC TTTCTTCAA 3780
 80 AGAGGCTCTG CCGGCTCCC TCCACCCAC TGCACCCATG AGACTCGGTC CAGAGTCCA 3840
 TTCCCCAGGT GGGAGCCAA TGTCAAGGAG GTCTTTCCA CCAACACTT TTAGCTGCT 3900
 GGGAGGTGAC CATAGGGCTC TGCTTTTAAA GATATGGCTG CTTCAAAGGC CAGAGTACA 3960
 GGAAGGACTT TCTCAGGGA GATTAGTGGT GATGGAGAGG AGAGTTAAA TGACCTCATG 4020
 TCTCTCTGT TCCAGGTTT GTTGAAGTTT CACTCTTCTA ATGCAAGGCT CTCACACTGT 4080
 GAACCACTTA GATGTGATC ACTTTCAGGT GGCAGGAAT GTTGAATGTC TTTGGCTCAG 4140
 TCTATTAAAA AAAGATACTT ATTTGAAAGT TCTCAGAGTT GTACATATGT TTCACAGTAC 4200

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

AGGATCTGTA CATAAAGIT TCTTTCCTAA ACCATTCAAC AAGAGCCAT ATCTAGGCAT 4260
TTTCTTGGA GCACAAATTT TCTTATTGCT TAGAAAATTG TCCTCCTGT TATTTCGTG 4320
TGTAAGACTT AAGTIGAGTTA GGTCTTTAAG GAAAGCAAG CTCCTCTGAA ATGCTTGTCT 4380
TTTTTCTGT GCGCAATATAG CTGGTCCCTT TCGGGAGTT AGATGTATAG AGTGTTTGTA 4440
TGTAACACTT TCTTGTAGGC ATCACCATGA ACAAAGATAT ATTTCTATIT TATTATATAT 4500
ATGTGCACCT CAAGAAGTCA CTGTCCAGAGA AATAAAGAT TGCTTTAAT GTCATGATTG 4560
GAGATGTCTT TTGCATTGCT TGGAAGGGGT GTACCTAGAG CCAAGGAAT TGGCTCTGGT 4620
TTGGAATAAT TTTGCTGTTA TTATAGTAAA CATACAAAGG ATGTCAAAAA AAAAAAATAA 4680
AAAAAATAA AAAAAAATAA AA

Seq ID NO: 142 Protein sequence
Protein Accession #: Eos sequence

1 11 21 31 41 51
MQAAGRQDFL FKAMLTISWL TLTCFPGATS TVAAGCPDQS PELQPNWPGH DQDREHVHIGQ 60
GKTLILLTSSA TVYSIHISEG GKLVIKDEDE FIVLRTRHIL IDNGGELHAG SALCPFGQNF 120
TILLYGRADE GIQDFPYGGL KYIGVGKGA LELHGGKLS WFLNKTLPF GMAEGGYFF 180
ERSWGRGVV VHVDPKSGT VIHSDRFDY RSKKESERLV QYLNAPDGR ILSVAVNDEG 240
SRMLDDMARK AMTKLGSKHF LHLGFRHFW FLTVKGNPSS SVEDHIEYHG HRGSAARVF 300
KLFQTBEGY FNVSLSSKRV QDVENTENFD HDKVSQYKGG EKISDLWKAH FGKICNRPID 360
IQATTMDGVN LSTEVVYKRG QDYRFACYDR GRACRSYKVR FLCGKPVPRK LTVTIDTNVN 420
STILNLEDNV QSWKPGDTLV IASTDYSMYQ AEEFQVLPCR SCAPNQVKA GKPMYLIHGE 480
EIDGDMRAE VGLLSRNTIV MGEMEDKCTP YRNHICNPF FDTFGSHKF ALGKAAHLE 540
GTLEKHMGGQ LVGQYPIHEH LAGDVDERGG YDPETYIRDL SIHETFSRCV TVHSGNGLLI 600
KDVVGNISLG HCFFTEDGPE ERNTEHCLG LLVKSGLLP SDRDSKMKCM ITEDSYPGYI 660
PKPRQCNV STFMANPMN NLINCAAGS EETGFNFIH HVTGFSVGM YSPGYSEHIP 720
LGFYNNRAH SNYRAGMTID NGVKTTEASA KDKRPLSII SARYSPHQDA DPLKPREPAI 780
IRHFIAKKNQ DHGANLRGGD VMLDSCHFRG EAQBPPLLTG MKAGGILLGG DEASGMAQGG 840
FSPPCRLKLV LVTGSPFAH VSLAHS

Seq ID NO: 143 NM DNA sequence
Nucleic Acid Accession #: NM_006017
Coding sequence: 38..2635

1 11 21 31 41 51
CCAAGTCTTA CTTCTATGTT GGAGGATCTT GCTAGCTATG GCGCTCGTAC TCGGCTCCCT 60
GTTGCTGCTG GCGCTGTGCG GGAACCTCCTT TTCAGGAGGG CAGCCTTCAT CCACAGATGC 120
TCTTAAGGCT TGGAATATAG AATGCTCCGC AACAAATAT GAGACCCAG ACTCCCATAA 180
AGCTGGACCC ATTGGCATTC TCTTTGAAC AGTCATATC TTTCTCTATG TGGTACAGCC 240
GCGTGATTTC CCAAGAGATA CTTTGAGAAA ATTCCTACAG AAGGCATATG AATCCAAAT 300
TGATTATGAC AGCCAGAAA CTGTAATCTT AGGTCTAAG ATTGTCTACT ATGAAGCAGG 360
GATTATCTA TGCCTGTGCC TGGGCTGTCT GTTTATTAAT CTGATGCGCT TGGTGGGGTA 420
TTTCTTTTGT ATGTGTCGTT GCTGTAAACA ATGTGGTGA GAAATGCACC AGCGACAGAA 480
GGAAATGGG CCGCTCCGGA GGAATGCTT TGCAATCTCC CTGTTGGTGA TTTGTATAAT 540
AATAAGCMTT GGCATCTTCT ATGTTTGTGT GSCAATCTAC CAGGTAGAA CCGGATCAA 600
AAGGAGTCGG AAATGCGCAG ATAGCAATTT CAAGGACTTG CGAATCTCT TGAATGAAC 660
TCCAGAGCAA ATCAATATA TATGGCCCA GTACAACTT ACCAAGGACA AGCGTTTAC 720
AGATCTGAAC AGTATCAATT CAGTGCTAGG AGGCGGAAT CTTGACCGAG TGAGACCCAA 780
CATCATCCCT GTTCTTGATG AGATTAGTC CATGGCAACA GCGATCAAG AGACCAAGA 840
GGCGTTGGAG AACATGAACA GCACCTTBA GAGCTTGCAC CAACAAAGTA CACAGCTAG 900
CAGCAGTCTG ACCAGCGTGA AAATAGCCT GCGGTCTCT CTCAATGACC CTCGTGCTT 960
GGTGATCTCA TCAAGTGAAA CCTGCAACAG CATCAGATT TCTCTAAGT AGCTGAATAG 1020
CAACCTGAAA CTGAGGCGAG TTCCACCGGT GGATGAGAAA CTTGACAAAG TTAATAACGT 1080
TCTTAGGACA GATTGGATG GCCTGGTCCA ACAGGCTAT CAATCCCTTA ATGATATACC 1140
TGACAGAGTA CAACGCCAAA CCACGACTGT CGTAGCAGGT ATCAAAGGG TTTGAATTC 1200
CATTGGTTCA GATATCGACA ATGTAATCA GCGTCTTCTT ATTCAGGATA TACTCTCAGC 1260
ATTCTCTGTT TATGTATAA ACACTGAAAG TTACATCCAC AGAANTTAC CTACATGGA 1320
AGAGTATGAT TCATCTGCTT GCGTGGGTGG CCTGCTCATC TGCTCTCTGC TGACCTCAT 1380
CGTGATTTTT TACTACCTGG GCTTACTGTG TGGCGTGTGC GGTATGACA GGCATGCCAC 1440
CCGACCCACC CGAGGCTGTG TCTCCACAC CGGAGGCGTC TTCTCATGG TTGGAGTTGG 1500
ATTAAGTTTC CTCTTTTGCT GATATTGAT GATCATGTG GTTCTTACCT TTGTCTTTGG 1560
TGCAAAATGG GAAAACTGA TCTGTGAACC TTACAGAGC AAGGAATAT TCCGGGTTTT 1620
GGATACACC TACTTACTAA ATGAAGACTG GGAATACTAT CTCTCTGGGA AGCTATTTAA 1680
TAAATCAAAA ATGAAGCTCA CTTTGAACA AGTTTACAGT GACTGCAAAA AAAATAGAGG 1740
CACTTACGGC ACTTCTCAC TGCAGAACAG CTTCAATATC AGTGAACATC TCAACATTAA 1800
TGAGCATACT GGAAGCATAA CAGTGAATT GGAAGTCTG AAGGTAAATC TTAATATCTT 1860
TCTGTTGGGT GCAGCAGGAA GAAAAACCT TCAGGATTTT GCTGCTTGTG GAATAGACAG 1920
AATGAATTTT GACAGCTACT TGGCTCAGC TGGTAAATCC CCGCAGGAG TGAATCTTTT 1980
ATCATTTGCA TATGATCTAG AAGCAAAAGC AAACAGTTTG CCCCAGGAA ATTTGAGGAA 2040
CTCCCTGAAA AGAGATGCAC AAATATTAA AACATTCAC CAGCAACGAG TCCTTCTAT 2100
AGAACATCA CTGAGCCTC TATACCRAAG CGTCAAGATA CTTCAAGCA CAGGGAATGG 2160
ATTGTTGGAG AGAGTAACTA GATTCTAGC TTCTCTGGAT TTTGCTCAGA ACTTCATCAC 2220
AAACATACT TCTCTGTGTA TTATTGAGGA AACTAAGAAG TATGGGAGAA CAATAATAGG 2280
ATATTTTGA AATATCTGC AGTGGATCGA GTTCTCTATC AGTGAGAAAG TGGCATCTGT 2340
CAAACTGTG GCCACCGCTC TAGATACGTC TGTGTATGTC TTTCTGTGTA GCTACATTAT 2400
CGAACCTTG AATTGTTTTT GGTTTGGCAT AGGAAAGCT ACTGTATTTT TACTTCGCGC 2460
TCTAATTTTT GCGGTAAACG TGGCTAAGTA CTATGTCGA ATGGATTCCG AGGACGTGTA 2520
CGATGATGTT GAAACTATAC CCAATGAAAA TATGAAAAAT GGTAAATAGT GTTATCATAA 2580
AGATCATGTA TATGATATTC ACAATCTGT TATGACAAGC CCAATCACAAT ATTGATAGT 2640
GATGTTGAAA CTGCTTGAGC ATCAGGATAC TCAAGTGGA AAGGATCACA GATTTTGGGT 2700
AGTTTCTGG TCTACAGGA CTTTCCAAAT CCAGGAGCAA GCGCAGTGGC AACGTAGTGA 2760
CTCAGGCGGG CACCAAGGCA ACGGCACCAT TGGTCTCTGG GTAGTGTCTT AAGAAATGAAC 2820

5	ACAATCACGT	TATAGTCCAT	GGTCCATCAC	TATTCAGGGA	TGACTCCCTC	CCTTCCCTGC	2880
	TATTTTGT	TTTACTTTT	TTACACTGAG	TTTCTATTTA	GACACTACAA	CATATGGGGT	2940
	GTTTGTCC	ATTTGGATGCA	TTTCTATCAA	AACCTATCA	AATGTGATGG	CTAGATTCTA	3000
	ACATATGGCC	ATGTGTGGAG	TGTGCTGAAC	ACACACCACT	TTACAGGAAA	GATGCATTIT	3060
	GTGTACAGTA	AACGGTGAT	ATACCTTTTG	TTACACAGAA	GTTTTTTAAA	CAAAAGAGTA	3120
	TTATAGGACT	TTCTTCTAAA	TGAGCTAAAT	AAGTCACCAT	TGACTTCTTG	GTGCTGTGGA	3180
	AAATAATCCA	TTTCACTAA	AAGTGTGTGA	AACCTACAGC	ATATTCTTCA	CGCAGAGATT	3240
	TTTCATCTATT	ATACITTTATC	AAAGATTGGC	CATGTTCCAC	TTGGAAATGG	CATGCAAAAG	3300
10	CCATCATAGA	GAAACCTGCG	TAACTCCATC	TGACAAATTC	AAAAGAGAGA	GAGAGATCTT	3360
	GAGAGAGAAA	TGCTGTTCTG	TCAAAAGTGG	AGTGTGTTTA	ACAGATGCCA	ATTACGGTGT	3420
	ACAGTTTAAC	AGAGTTTICT	GTTCGATTAG	GATAAACATT	AATTGAGATG	CAGCTAACAT	3480
	GAGTATCATC	AGACTAGTAT	CAAGTGTTCT	AAAATGAAAT	ATGAGAAGAT	CCTGTACAAA	3540
	TTCTTAGATC	TGGTGTCAG	CATGGATGAA	ACCTTGAGT	TTGGTCCCTA	AATTGTCATG	3600
15	AAAGCAACAG	GTAAATATTC	ATTTGCTTCA	GGAGTTTCAT	GTGGATCTG	TCATTATCAA	3660
	AAGTGTACAG	CAATGAGGAA	CTGGTCCGAC	AAAATTTAAC	GTGATGTAA	TGGAATTCCA	3720
	GATGTAGGCA	TTCCCCCCAG	GTCITTTTCAT	GTGCAGATTG	CAGTTCTGAT	TCATTTGAAT	3780
	AAAAAGGAAC	TTGG					

Seq ID NO: 144 Protein sequence
Protein Accession #: NP_006008.1

25	1	11	21	31	41	51	
	MALVLGSLLL	LGLCGNSFSG	GQPESTDAEK	AMNVELPATN	YETQDSHKAG	PIGILFELVH	60
	IFLYVQPRD	FPEDTLRKEL	QKAYESKIDY	DKPETVILGL	KIVYYEAGII	LCCVLGLDFI	120
	ILMPLVGYYF	CMCRCCNKG	GEMEQRKEN	GFPLRCKPAI	SLIVICIIIS	IGIFYGFVAN	180
	HQVTRIKES	RKLADSNFED	LRTLLNETPE	QIKYILAQYN	TTKDKAFIDL	NSINSVLGGG	240
	ILDRLENII	PVLDEIKSMA	TAIKETKEAL	ENMNSTLKSL	HQOSTQLSSS	ITSVKTSLRS	300
30	SLMDPLCLVH	PSSETCSNR	LBSLQINSNF	ELRQLFPVDA	ELDNVNNVLH	TDLDGLVQQG	360
	YQSLNDIPDR	VQRQITTVVA	GIKRVLSNG	SDIDNVTRL	PIQDLSAFS	VYVNTESYI	420
	HEMLPLEEY	DSYWNLGGIV	ICSLTLTIV	FYYLGLLGGV	CGYDRHATPT	TRGCVSNTGG	480
	VFLMVGVLG	FLFCNLIMII	VVLTFVFGAN	VEKLICFPYT	SKELFRVLDI	PYLINEDWEY	540
	YLSGKLENKS	KMKITFEQVY	SDCKKNRGTY	GTLLHLQNSFN	ISEHLNINEH	TGSISSELES	600
35	LKVNLTFL	GAAGRNLDQ	FAACGIDRMN	YDSYLAQTGK	SPAGVNLLSF	AYDLEAKANS	660
	LPPGNLNSL	KRDAQTIKTI	HQRVLPIEQ	SLSTLYQSVK	ILQRTGNGLL	ERVTRILASL	720
	DFAQNFITNN	TSSVILEETK	KYGRITIGYF	EHYLQWIEFS	ISEKVASCKP	VATALDTAVD	780
	VFLCSYIIDP	LNLFWFGIGK	ATVFLLPALI	FAVELAKYFR	RMSSEIDVDD	VETITPMNME	840
	NGNNGYKDH	VYGINPVMIT	SFSQH				

Seq ID NO: 145 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 38..2605

45	1	11	21	31	41	51	
	CCAAGTCTA	CCTCATGTTT	GGAGGATCTT	GCTAGCTATG	GCCCTCGTAC	TOGGCTCCCT	60
	GTGTCTCTG	GGGCTGTGGG	GGAACCTCCT	TTGAGGAGGG	CAGCCCTTCAT	CCACAGATGC	120
	TCCTAAGGCT	TGGAATTATG	AATTGCCTGC	AACAAATAT	GAGACCCCAAG	ACTCCCATAA	180
50	AGCTGGACCC	ATTGGCATTTC	TCTTTGAACT	AGTGCATATC	TTTCTCATG	TGATCAGCCC	240
	GCGTATTTTC	CCAGAAGATA	CTTTGAGAAA	ATTCTTACAG	AAGGCATATG	AATCCAAAT	300
	TGATTATGAC	ATTGTCTACT	ATGAAGCAGG	GATTATTCTA	TGCTGTGTCC	TGGGCTGTCT	360
	GTTTATTATT	CIGATGCCCT	TGGTGGGGTA	TTTCTTTTGT	ATGTGTCTGT	GCTGTAAACAA	420
	ATGTGGTGGA	GAAATGCACC	AGCGACAGAA	GGAATAATGG	CCCTTCTCTG	GGAAATGCTT	480
55	TGCAATCTCC	CTGTTGGTGA	TTTGTATAAT	AATAAGCATT	GGCATCTTCT	ATGGTTTGT	540
	GGCAATCAC	CAGGTAAAGAA	CCCGGATCAA	AAGGATGCG	AACTGGCAG	ATAGCAATTT	600
	CAAGGACTTG	CGAATCTCT	TGAATGAAC	TCCAGAGCAA	ATCAATATA	TATTGGCCCA	660
	GTACCAACACT	ACCAGGACAA	AGCGGTTTAC	AGATCTGAAC	AGTATCAATT	CAGTGTAGG	720
	AGGCGGAATT	CTGACCCGAC	TGAGACCCAA	CATCATCCCT	GTCTCTGATG	AGATTAAATC	780
60	CATGGCAACA	GATCTCAAGG	AGACCAAGAA	GGCGTTGGAG	AACATGAACA	GCACTTGA	840
	GAGCTTGAC	CAACAAGTA	CACAGCTTAS	CAGCAGTCTG	ACCAGCGTGA	AACTAGCCT	900
	GCGTCTATCT	CTCAATGACC	CTCTGTGCTT	GGTGATCCCA	TCAAGTGAA	CCTGCAACAG	960
	CATCAGATTG	TCTCTAAGCC	AGCTGAATAG	CAACCCTGAA	CTGAGGCAGC	TTCCACCCGT	1020
	GGATGCGA	CTTGACAAAG	TTAATAAGCT	TCTTAGGACA	GATTTGGATG	GCTGTGTCCA	1080
65	ACAGGGCIAT	CAATCCCTTA	ATGATATACC	TGACAGAGTA	CAACGCCAAA	CCACGACTGT	1140
	CGTAGCAGGT	ATCAAAAGGG	TCTTGAATTC	CATTGGTICA	GATATCGACA	ATGTAATCTA	1200
	GCTCTTCT	ATTCCAGATA	TACTCTCAGC	ATTCTCTGTT	TATGTTAATA	AACTGAAG	1260
	TTACATCCAC	AGAAATTTAC	CTACATTGGA	AGAGTATGAT	TCATACTGGT	GGCTGGGTGG	1320
	CCTGGTCATC	TGCTCTCTGC	TGACCTCAT	CGTATTFTT	TACTACTGCG	GCTTACTGTG	1380
70	TGGCGTGTG	GGCTATGACA	GGCATGCCAC	CCCGACCACC	CBAGGCTGTG	TCTCCACAC	1440
	CGGAGGCGTC	TTCTCATGG	TGGAGTGG	ATTAAGTTTC	CTCTTTTGTCT	GGATATTGAT	1500
	GATCATTGTG	GTTCTTACCT	TGTCTTTGG	TGCAATGTG	GAAAACTGA	TCTGTGAACC	1560
	TTACACAGAC	AAGGAATTAT	TCCGGTTTT	GGATACACCC	TACTTACTAA	ATGAAGACTG	1620
	GGATACATAT	CTCTCTGGGA	AGCTATTATA	TAAATCAAAA	ATGAAGCTCA	CTTTGAAACA	1680
75	AGTTTACAT	GATCTCAAAA	AAAATAGAGG	CACITACGGC	ACTCTTCAAC	TGCAGAACAG	1740
	CTTCAATATC	AGTGAACATC	TCAACATTAA	TGAGCATACT	GGAAGCATAA	GCACTGAATT	1800
	GGAAAGTCTG	AAGGTAATC	TTAATATCTT	TCTGTTGGGT	GCAGCAGGAA	GAATAAACCT	1860
	TCAGGATTTT	CTGTCTTGTG	GAATAGACAG	AATGAATTAT	GACAGCTACT	TGGCTCAGAC	1920
	TGGTAATATC	CCCGCAGGAG	TGAATCTTTT	ATCATTTGCA	TATGATCTAG	AAGCAAAAGC	1980
80	AAACAGTTTG	CCCTCAGGAA	ATTTGAGGAA	CTCCTTGAAA	AGAGATGCAC	AAACTATTAA	2040
	AACATTCAC	CAGCAACGAG	TCTTCTCTAT	AGAACATCA	CTGAGCCTCT	TATACCAAG	2100
	CGTCAAGATA	CTTCAACGCA	CAGGGAATGG	ATTGTTGGAG	AGAGTACTTA	GGATTCTAGC	2160
	TTCTCTGGAT	TTTGTCTCAG	ACTTCTCTAC	AAACATTAAT	TCTCTGTGTA	TTATTGAGGA	2220
	AACTAAGAG	TATGGGAGAA	CAATAATAGG	ATATTTTGAA	CATTATCTGC	AGTGGATCGA	2280
	GTTCTCTATC	AGTGAGAAAG	TGGCATCGTG	CAACCTGTG	GCCACCGCTC	TAGATACTGC	2340

5
 10
 15
 20
 25

```

TGTTGATGTC TTTCTGTGTA GCTACATTAT CGACCCCTTG AATTGTGTTT GGTTGGCAT 2400
AGGAAAAGCT ACTGTATTTT TACTTCCGSC TCTAATTTTT GCGGTAAAAC TGGCTAAGTA 2460
CTATCGTCGA ATGGATTCGG AGGACGCTGA CGATGATGTT GRAACTATAC CCATGAAAAA 2520
TATGGAAAAAT GGTAAATAAT GTTATCATAA AGATCATGTA TATGGTATTC ACAATCCGT 2580
TATGACAAGC CCATCACAAC ATTGATAGCT GATGTTGAAA CTGCTTGAGC ATCAGGATAC 2640
TCRAAGTGGA AAGGATCACA GATTTTTCGT AGTTTCTGGG TCTACAAGGA CTTTCCAAAT 2700
CCAGGAGCAA CGCCAGTGGC AACGTAGTGA CTCAGGCGGG CACCAAGGCA ACGGCACCAT 2760
TGGTCTCTGG GTAGTGCTTT AAGAATGAAC ACAATCAGCT TATAGTCCAT GGTCCATCAC 2820
TATTCAGGA TGACTCCCTC CTTCTCTGTC TATTTTGTGTT TTTTACTTTT TTACACTGAG 2880
TTTCTATTTA GACACTACAA CATATGGGGT GTTTGTTCOC ATTGGATGCA TTTCTATCAA 2940
AACTCTATCA AATGTGATGG CTAGATTCTA ACATATTCOC ATGTGTGGAG TGTGCTGAAC 3000
ACACACACAT TTACAGGAAA GATGCATTTT GTGTACAGTA AACGGTGTAT ATACCTTTTG 3060
TTACCACAGA GTTTTTTAAA CAAATGAGTA TTATAGGACT TTCTTCTAAA TGAGCTAAAT 3120
AAGTCACCAT TGACTTCTTG GTGCTGTGTA AAATAATCCA TTTTCACTAA AAGTGTGTGA 3180
AACCTACAGC ATATTCTTCA CGCAGAGATT TTCTCTATT ATACTTTATC AAAGATTGGC 3240
CATGTTCCAC TTGAAAATGG CATGCAAAAG CCATCATAGA GAAACCTGCG TAACCTCCATC 3300
TGACAAATTC AAGAGAGAGA GAGAGATCTT GAGAGAGAAA TGCTGTTGCT TCAAAAGTGG 3360
AGTGTGTTTA ACAGATGCCA ATTACGGTGT ACAGTTTAA CAGAGTTTCT GTTGCAATAG 3420
GATAAACATT AATTGGATGG CAGCTAACAT GAGTATCATC AGACTAGTAT CAGGTGTCTT 3480
AAAATGAAT ATGAGAAGAT CCTGTACAA TTCTTAGATC TGGTBTCCAG CATGGATGAA 3540
ACCTTTGAGT TTGTTCCCTA AATTTCGATG AAAGCACAAG GTAAATATTC ATTTGCTTCA 3600
GGAGTTTCAT GTTGGATCTG TCTTTATCAA AAGTGATCAG CAATGAAGAA CTGGTCGGAC 3660
AAAATTTAAC GTTGATGTAA TGAATTCCA GATGTAGGCA TTCCCCCAG GTCTTTTCAT 3720
GTGCAGATTG CAGTTCTGAT TCATTTGAAT AAAAAAGAAC TTGG
  
```

Seq ID NO: 146 Protein sequence
 Protein Accession #:

30
 35
 40
 45

```

1 11 21 31 41 51
MALVIGSLLL LGLCQNSPSG GQPSSTDAPK AMNYELPATN YETQDSHKAG PIGILFELVH 60
IFLYVVQPRD FPDATLRKFL QKAYESKIDY DIVYYRAGII LCCVLGLLFI ILMPVGVYFF 120
CMCRCCNKG GEMHQKQEN GPFLRKCFAI SLIVICIIIS IGFYGFVAN HQVRTRIRS 180
RLKLDSENFQ LRTLLNETPE QIKYILAQYN TTKDKAFTDL NSINSVLGGG ILDRLRPMII 240
FVLDEIKSMA TAIKTKKRAL EMMNSTLKSLL HQQSTQLSSS LSVKTSLSRS SLNDPLCLVH 300
PSSETONSIR LSLSQLNSNP ELRLPLPVDA ELDNVNVLRL TDLDGLVQQG YQSLNDIPDR 360
VQRQTITVA GIKRVLSIG SDIDNVTRL PIQDILSAPS VYVNNTESYI ERNLPTLEBY 420
DSYWNLGGV ICSLLTLVIV FYLLSLGCGV CGYDRHATPT TRGCVSTGG VFLMVGVGLS 480
FLFCWILMII VVLTFFVGAN VEKLICEFYT SKELFRVLDL PYLLNEDMEY YLSGKLFNKS 540
EMKLTPEQVY SDCKMRGIY GFLHLQNSFN ISEHLNINEH TGSISSELES LKVNLIPIEL 600
GAAGRKLQD FARCGIDRMN YDSYLAQTGK SPAGVNLLSF AYDLEAKANS LPPGNLRNSL 660
KRDAQTILKT HQQRVLPLEQ SLSTLYQSVK ILQRTGNGLL ERVTRILASL DPAQNFITNW 720
TSSVIIIEBK KYGRITIGYF EHYLQWIEFS ISEKVASCKE VATALDTAVD VFLCSYIIDP 780
LNLFWFGIGK ATVFLPALI FAVKLAKTYR RMDSEVDYDD VETIPMKMWE NGNNGYHRDH 840
VYGIHNPVMI SPSSQH
  
```

Seq ID NO: 147 DNA sequence
 Nucleic Acid Accession #: BC012089.1
 Coding sequence: 236..2806

50
 55
 60
 65
 70
 75
 80

```

1 11 21 31 41 51
ATGCTCTCAG CTCTCCCGCC CGCGGATGGT GCCTGAGTGT AATGACCCCC TTGGAGAAC 60
TTCTTCGSCA TCCCTCGCCT CAGGCCAGCC TCAGACAGAA AACTGAAGAT TCAGCAGATC 120
CAGTGTCTCC TGCTCTCTCT CTGCCAGGGA ACAGGCTTGC CTTCGCCAAG GCTTCCAGAA 180
GCTCTGAGGC AGGAGGCACC AAGTCTTACC TCATGTTTGG AGGATCTTGC TAGCTATGGC 240
CCTCGTACTC GGCTCCCTGT TGCTGCTGGG GCTGTGCGGG AACTCTTTT CAGGAGGGCA 300
GCCTTCATCC ACAGATGCTC CTAAGGCTTG GAATTATGAA TTGCCCTGCA CAAATTATGA 360
GACCCAGAGC TCCCATTAAG CTGAGCCCAT TGGCATCTCT TTGAACTAG TGCATATCTT 420
TCTCTATGTG GTACAGCCGC GTGATTTCCT AGAAGATACI TTGAGAAAT TCTTACAGAA 480
GGCATATGAA TCCAAAATTG ATTATGACAA GATTGTCTAC TATGAAGCAG GGATTATCT 540
ATGCTGTGTC CTGGGCTGTC TGTTTATAT TCTGATGCT CTGGTGGGAT ATTTCTTTTG 600
TATGTGTGCT TGCTGTAAAC AATGTGTGAG AGAAATGCAC CAGCGACAGA AGGAAAATGG 660
GCCCTTCTCT GAGCAATGCT TTGCAATCTC CCTGTTGGTG ATTTGTATAA TAATAAGCAT 720
TGGCATCTTC TATGTTTGTG TGGCAATCA CCAGGTAAAG ACCCGATCA AAAGGAGTCG 780
GAAACTGGCA GATAGCAAT TCAAGGACTT CGCACTCTC TTGAATGAA CTCCAGAGCA 840
AATCAAAAT ATATTGGCCC AGTACAAAC TACCAAGGAC AAGGCGTTCA CAGATCTGAA 900
CAGTATCAAT TCAGTGCTAG GAGGCGGAAT TCTTGACCGA CTGAGACCCA ACATCATCCC 960
TGTCTTGAT GAGATTAAAT CCATGGCAAC AGCGATCAAG GAGACCAAG AGGDSITGGA 1020
GAACATGAAC AGCACTTGA AGAGCTTGCA CCAACAAAGT ACACAGCTTA GCAGCAGTCT 1080
GACCAGCGTG AAAACTAGCC TGCGGTATC TCTCAATGAC CCTCTGTGCT TGGTGATCC 1140
ATCAGATGAA ACCTGCAACA GCATCAGATT GTCTCTAAGC CAGCTGAATA GCAACCCGTA 1200
ACTGAGGCGG GTTCCACCCG TGATGCGAGA ACTTGCAAC GTTAATAACG TTCTTAGGAC 1260
AGATTGGAT GGCCTGTGTC AACAGGGCTA TCAATCCCTT AATGATATAC CTGACAGAGT 1320
ACAAAGCCAA ACCAAGACTG TGTAGCAGG TATCAAAGG GTCTTGAAAT CCATTGGATC 1380
AGATATGAC AATGTAACTC AGGTCCTCC TATTCAAGAT ATACTCTCAG CATTCCTGAT 1440
TTATGTTAAT AACACTGAAA GTTACATCCA CAGAAATTTA CCTACATTGG AAGAGTATGA 1500
TTCATACTGG TGGCTGGGTG GCCTGGTCT CATGCTCTCT CTGACCCCTA TGTGATTTT 1560
TTACTACTCT GGCCTACTGT GTGCGGTGTG CGGCTATGAC AGGCATGCCA CCCGACCCAC 1620
CGAGGCGTGT GTCTCCAACA CCGAGGCGCT CTTCTCTATG GTTGGAGTTG GATTAAAGTT 1680
CCTCTTTTGC TGGATATTGA TGATCAITGT GGTTCCTACC TTGTCTTTG GTGCAAATGT 1740
GGAAAACTG ATCTGTGAAC CTTACAGAG CAAGGATTTA TTCCGGGTTT TGGATACACC 1800
CTACTTACTA AATGAGACT GGGGAATACTA TCTCTCTGGG AAGCTATTTA ATAAATCAAA 1860
AATGAAGCTC ACTTTTGAAC AAGTTTACAG TGACTGCAAA AAAAATAGAG GCACCTACGG 1920
  
```

5
10
15
20
25
30
35

```

CACICTTCAC CTGCAGAAC GCTTCAATAT CAGTGAACAT CTCAACATTA ATGAGCATA 1980
TGGAAAGCATA AGCAGTGAAT TGGAAAGTCT GAAGGTAAAT CTTAATATCT TTCTGTTGGG 2040
TGCAGCAGGA AGAAAAACC TTCAGGATTT TGCTGCTTGT GGAATAGACA GAATGAATTA 2100
TGACAGCTAC TTGGCTCAGA CTGGTAAATC CCCCSCAGGA GTGAATCTTT TATCATTTGC 2160
ATATGATCTA GAAGCAAAAG CAAACAGTTT GCCCCAGGA AATTGAGGA ACTCCCTGAA 2220
AAGAGATGCA CAAACTATTA AAACAATTCA CCAGCAACGA GTCCCTCCCTA TAGAACATC 2280
ACTGAGCACT CTATACCAA GCGTCAAGAT ACTTCAACGC ACAGGGAATG GATTGTTGGA 2340
GAGAGTAACT AGGATTCCTAG CTCTCTGGA TTTGCTCAG AACTTCATCA CAAACAATAC 2400
TTCCCTCTGTT ATTATGAGG AAACATAAGAA GTATGGGAGA ACAATAATAG GATATTTTGA 2460
ACATTATCTG CAGTGGATG AGTTCTCTAT CAGTGAGAAA GTGGCATCGT GCAAACTGT 2520
GGCCACCGCT CTAGTACTG CTGTTGATGT CTTCTGTGT AGCTACATTA TCGACCCCTT 2580
GAATTTGTTT TGGTTGCGA TAGGAAAAGC TACIGTATTT TTACTTCCGG CTCTAATTTT 2640
TGCGGTAAAA CTGGCTAAGT ACTATCGTCG AATGGATTCG GAGGACGTTG ACBATGATGT 2700
TGAAACTATA CCACTGAAAA ATATGGAATA TGGTAATAAT GATTATCATA AAGATCATGT 2760
ATATGGTATT OCCATCCCTG TTATGACAAO OCCATCACA CATTGATAGC TGATGTTGAA 2820
ACTGCTTGAG CATCAGGATA CTCAGGATGG AAAGGATCAC AGATTTTTCG TAGTTTCTGC 2880
GTCTACAGAG ACTTACCAA TCCAGGAGCA ACGCCAGTGG CAACGTAGTG ACTCAGGCGG 2940
GCACCAAGGC AACGGCACCA TTGGTCTCTG GGTAGTGCCT TAAGAATGAA CACAATCAG 3000
TTATATCCCA TGGTCCATCA CTATTCAGG ATGACTCCCT CCCTTCCCTG CTATTTTGT 3060
TTTTTACTTT TTATCATGA GTTCTATTT AGACACTACA ACATATGGGG TGTGTGTTCC 3120
CATTTGATGC ATTTCTATCA AAATCTATC AAATGTGATG GCTAGATTTT AACATATGTC 3180
CATGTGTGGA TGGTTGCGA CACACACCAG TTTACAGGAA AGATGCATTT TGTGTACAGT 3240
AAACGGTGTA TATACCTTTT GTTACCACAG AGTTTITTA ACRAATGAGT ATTATAGGAC 3300
TTTCTCTTAA ATGAGCTAAA TAAGTCACCA TTGACTTCTT GGTGCTGTTG AAAATAATCC 3360
ATTTTCACTA AAAGTGTGTG AAACCTACAG CATATTCTTC ACGCAGAGAT TTCTATCTAT 3420
TATACTTTAT CAAAGATTTG CCATGTTCCA CTGGAATG GCATGCAGAA GCCATCATAG 3480
AGAAACCTGC GTATTCTCT CAGACAAATT CAAAAGAGAG AGAGAGATCT TGAGAGAGAA 3540
ATGCTGTTG TTCAAAGTG GAGTTGTTT AACAGATGCC AATTACGGTG TACAGTTTAA 3600
CAGAGTTTTC TGTGCTATTA GGATAAACAT TAATTGAGT GCAGCTAACA TGAGTATCAT 3660
CAGACTAGTA TCAAGTGTTC TAAATGAAA TATGAGAAGA TCCTGTGACA ATTCTTAGAT 3720
CTGGTGTTCA GCATGATGA AACCTTTGAG TTTGGTCCCT AAATTTCAT GAAAGCACAA 3780
GGTAAATATT CATTTGCTTC AGGAGTTTCA TGTGGATCT GTCAATTATCA AAAGTATCA 3840
GCAATGAAGA ACTGGTCGGA CAAAATTAA CGTGTATGTA ATGGATTC CCAGATGAGGC 3900
ATTOCCCCCA GGTCTTTTCA TGTGCAGATT GCAGTTCTGA TTCATTGAA TAAAAGGAA 3960
CTTGGAAAAA AAAAAAAA AAA

```

Seq ID NO: 148 Protein sequence
Protein Accession #: AAH12089.1

40
45
50
55

```

1 11 21 31 41 51
MALVLGSLLL LGLCGNSFSQ QPSSSTDAEK AWNYELPATN YETQDSHKAG FIGILFELVH 60
IFLYVVQPRD FPDRTLRLKPL QKAYESKIDY DKIVYYEAGI ILCCVLGLLF IILMLVGYF 120
PCMCRCKNCK GGEMEQKQKE NGFFLRKCPA ISLVICIII SIGIFYGFVA NBQVTRIKR 180
SRKLADSNFK DLRLLNETP BQIKYILAQY NTKDKAFTD LNSINSVLGG GILDRLRPMI 240
IPVLDEIKSM ATAIEKTEEA LEMNNTLSKS LHQQSTQLSS SLTSVKTSLR SSINDPLCLV 300
HPSSETCNIS RLSLSLQNSN PELRLQPPVD AELDNVNNVL RTDLGLVQQ GYQSLNDIFD 360
RVQRGTTTVV AGIKRWLNSI GSDIDNVTOR LPIQDILSAF SVYVNNTEBY IHERNPTLKE 420
YDSYWLGLGL VICSLLLTIV IFYYLLELLE VCGYDRHATP TTRGCVSNTG GVFLMVGVL 480
SFLPCWILMI IUVLTFVFGA HVEKLICEPY TSKELFRVLD TPYLLNEDWE YLDSKLFNK 540
SKMKLTFFQV YSDCKNMRGT YGTLHLQNSF NISEHLNINE HTGSISSBLE SLKVNLIIFL 600
LGAAGRKNLQ DPAACGKIRM NYDSYLAQGT KSPAGVNLIS EAYDLBAKAN SLPPCNLRNS 660
LKRDAQTETK IEDQRLPIE QSLSTLYQSV KILQRTGMLG LERVTRILAS LDFAQNFITN 720
NTSSVRIIBET KKXGRITIGY FEHYLQNIET SISEKVASCK PVATALDTAV DVFLCSYIID 780
PLNLFWFQIG KATVFLPAL IPAVKLAKY RMDSEDEVYD DVSTIFMKNM ENGNNGYHKD 840
HVVYGHNFVM TSPSQH

```

Seq ID NO: 149 DNA sequence
Nucleic Acid Accession #: NM_033049.1
Coding sequence: 28..1566

60
65
70
75
80

```

1 11 21 31 41 51
CCACGCGTCC GAGCAAGAAC AGCTAAATG AAAGCCATCA TTCTCTTAC TCTTCTTCT 60
CTCCTTTCTG TAAACACAGC CACCAACCAA GGCAACTCAG CTGATGCTGT AACCAACCA 120
GAAACTCGGA CTAGTGTGTC TACAGTAGCT GCAGCTGATA CCCTGAAAC TAATTTCCCT 180
GAAACTGCTA GCACACAGC AAATACACCT TCTTCCCAA CAGCTACTTC ACCCTGCTCC 240
CCCATAAITA GTACACATAG TTCTCCACA ATTCCTACAC CTGCTCCCC CATATTAGT 300
ACACATAGTT CTTCCACAT TCTATACCT ACTGCTGCAG ACAGTGAATC AACTCAAA 360
GTAAATTCAT TAGCTACCTC TGACATAATC ACCGCTTCAT CTCCAAATGA TGGATTAATC 420
ACAATGGTTC CTCTGAAAC ACAAGTAAC AATGAAATGT CCCCACACAC AGAGACAA 480
CAATCATCAG GGCTCCGCTC TGGCAACGCT TTATGGAGA CCAGCACCTC AACACGACA 540
GGTCCGACGA ATCTTGCCTA AGATGATCCC TGTGCAGATA ATTGCTTATG TGTAAAGCTG 600
CATAAACAA GTTTTTCGCT GTGTTAGAA GGGTATTACT ACAACTCTTC TACATGTAG 660
AAAGGAAGG TATTCCTCGG GAAGATTTCA GTGACAGTAT CAGAAACATT TGACCCAGAA 720
GAGAAACATT CCATGGCCTA TCAAGACTTG CATAGTGAAT TTACTAGCTT GTTTAAAGAT 780
GTATTTGCA CATCTGTTA TGGACAGACT GTAATCTTA CTGTAAGCAC ATCTCTGTCA 840
CCAAGATCTG AAATGCGTGC TGATGACAAG TTTGTTAATG TAACAATAGT AACATTTT 900
GCAGAAACCA CAGGTGACAA TGAGAAGACT GTGACTGAGA AAATTAATAA AGCAATTAGA 960
AGTAGCTCAA GCAACTTTCT AAACATATGAT TTGACCCCTC GGTGTGATTA TTATGGCTGT 1020
AACGAGACTG CGGATGACTG CCTCAATGGT TTAGCATGCG ATTCGAAATC TGACCTGCAA 1080
AGGCTTAACC CACAGAGCCC TTTCTGCGTT GCTTCCAGTC TCAAGTSTCC TGATGCTGCT 1140
AACGACAGC ACAAGCAATG CTTAATAAAG AAGAGTGGTG GGGCCCTCGA GGTGCGTGC 1200
GTGCCCGCT ACCAGGAAGA TGCTAATGGG AACTGCCAAA AGTGTGCATT TGGCTACAGT 1260

```

5
10
15
20
25
30

```

GGACTCGACT GTAAGGACAA ATTTGAGCTG ATCCTCACTA TTGTGGGCAC CATCGCTGGC 1320
ATTGTGATTC TCAGCATGAT AATTGCATIG ATTGTGACAG CAAGATCAAA TAACAAACAG 1380
AAGCATATTG AAGAGAGAGAA CTGTATTGAC GAAGACTTTC AAAATCTAAA ACTGCGGTG 1440
ACAGGCTTCA CCAATCTTGG AGCAGAAGGG AGCGTCTTTC CTAAGGTGAG GATAACGGCC 1500
TCCAGAGACA GCCAGATGCA AAATCCCTAT TCAAGACACA GCAGCATGCC CCGCCCTGAC 1560
TATTAGAAATC ATAAGAAATGT GGAACCCGCC ATGGCCCCCA ACCAATGTAT AAGCTATTAT 1620
TTAGAGTGTG TAGAAGAGACT GATGGAGAAG TGAGCACCCAG TAAAGATCTG GCCTCCGGGG 1680
TTTTTCTTCC ATCTGACATC TGCCAGGCTC TCTGAATGGA AGTTGTGAAT GTTTGCAACG 1740
AATCCAGCTC ACITGCTAAA TAAGAATCTA TGACATTAATA TGTAGTAGAT GCTATTAGCG 1800
CTTGTGAGAG AGSTGGTTTT CTTCATCAG TACAAGTAC TGAGACAATG GTTAGGGTTG 1860
TTTTCTTAAT TCITTTCTCTG GTAGGGCAAC AAGAACCATT TCCAATCTAG AGGAAAGCTC 1920
CCAGCATATG CTTCCTCCTG GGCAAACATT GCTCTTGAGT TAAGTGAOCT AATTCCTCTG 1980
GGAGACATAC GCATCAACTG TGGAGGTCCG AGGGGATGAG AAGGGATACC CACCACCTTT 2040
CAAGGGTCAC AAGCTCACTC TCTGACAAAT CAGAATAGGG ACACTGCTTC TATCCCTCCA 2100
ATGGAGAGAT TCTGGCAACC TTTGAACAGC CCAGAGCTTG CAACCTAGCC TCACCCAAAG 2160
AGACTGGAAA GAGACATATC TCTCAGCTTT TTCAGGAGGC GTGCCCTGGG ATCCAGGAAC 2220
TTTTTGATGC TAATTAGAG GCTTGGACTA AAAATGTCCA CTATGGGGTG CACTCTACAG 2280
TTTTTGAAAT GCTAGAGAGC AGAAGGGGCA GAGAGTAAAA AACATGACCT GGTAGAAGGA 2340
AGAGAGTCAA AGGAAACTCG GTGGGGAGGA TCAATTAGAG AGGAGGCACC TGGGATCCAC 2400
CTTCTTCTT AGSTCCCTTC CTCCATCAGC AAAGGAGCAC TTCTCTAATC ATGCCCTCCC 2460
GAAGACTGGC TGGGAGAAGG TTTAAAAACA AAAATCCAG GAGTAAGAGC CTTAGGTGAG 2520
TTTTGAATTC GAGACAAACT GTCTGGCAAA GGGTGGGAGA GGGAGCTTGT CACTCAGGAGT 2580
CCAGCCCTCC AGCCTCGGGG TGTAGGTTCG TGAGGTGTGC CATTGGGGCC TCAGCTTCTC 2640
CTGGTGACAC DVGTSVYGG TVILTVSTSL SPRSEMRADD KFNVTIVTI LAETSDNEK 2700
ACAAATGGGG GCACCCACAT CCAGTACAG CTCTTACAAA TGTATTAGT GTCTTTTTC 2760
ATTTCTAAGC CCTTGTCTTC TTAAGAGTTA TTTTATTGT TATTATTAT TGTCTTGGAC 2820
TGTTAATTGT GAATGGTAAT GCATTAAGT GCCTTGTGTA GATGGTGAAA AAAAAAATAA 2880
AAAAAA

```

Seq ID NO: 150 Protein sequence
Protein Accession #: NP_149038.1

35
40
45

```

1 11 21 31 41 51
MKAIHLLELL ALLESVNTATN QMSADAVTT TETATSGPTV AAADTTETNF PETASTTANT 60
PSFPATSPAP PPIISTHSSS TIPTPAPPII STHSSSTIPI PTAADSEST NVNSLATSDI 120
ITASSPNDGL ITMVPBETQS MNEMSPTTED NQSSGPTTGT ALLETSTLNS TGPBNPCQDD 180
PCADNSLCVK LHMTSPCLCL EGYVYNSSTC KKGKVFPGKI SVTVSEYDFP BEKHSMAYQD 240
LHSEITSLFK DVGTSVYGG TVILTVSTSL SPRSEMRADD KFNVTIVTI LAETSDNEK 300
TVTEKINKAI RSHSSNELYN DLTLRCDYVG CMTADDCLN GLACDCKSDL QRPNPQSPFC 360
VASSLKCPDA CNAQHKQCLI KXSGGAPRCA CVFVQYEDAN GNCQKCAFYG SGLDCKDKFPQ 420
LILITVGTIA GIVILSMITIA LIVTARSNNK TKRIERENLI DEDFQNLKLR STGFTNLGAE 480
GSVPFKVRII ASRDSQMOPN YSRHSSMPRP DY

```

Seq ID NO: 151 DNA sequence
Nucleic Acid Accession #: Bos sequence
Coding sequence: 87..914

50
55
60
65
70
75
80

```

1 11 21 31 41 51
AGTACGCGGG GCGCGGTTTT GCCTCCGCGG CAGCTCTGGG CTCTTCTCAG CTGCGCGAGC 60
AGCTGCTCCA ATGCCCCGGA GTGGCCATGG GCGCCCCGCA CTGGTGGGAC CAGCTGCGAG 120
CTGTAGCTAC GAGAGTGGAC TGGTGCGAGG ACMACTACAC CATCGTGCTT GCTATOGCGG 180
AGTTCTACAA CAGATCAGC AATGTCTTAT TTTTCATTTT ACGGCCCMTC TGCAATGAGT 240
TGTTTTCGTA GTATGCAACA TGCTTCAACA GTGGCATCTA CTAAATCTGG ACTCTTTTGG 300
TGTAGTGGG AATTGGATCC GTCTACTTCC ATGCAACCCCT TAGTTTCTTG GGTGAGATGC 360
TTGATGAAGT TGCAGTCCCT TGGGTTCTGA TGTGTGCTTT GGCATGTGGT TCCCCAGAA 420
GGTATCTACC AAAGATCTTT CGGAATGACC GGGGTAGGTT CAAGGTGGTG GTGAGTGTCC 480
TGTCTGCGGT TACGACGTGC CTGGCATTTG TCAAGCCTCG CATCAACAAC ATCTCTCTGA 540
TGACCTTGGG AGTTCCTTGC ACTGCACTGC TCATCGCAGA GCTAAAGAGG TGTGACAAAC 600
TGCGTGTGTT TAAGCTGGGC CTCTTCTCGG GCCTCTGGTG GACCTCTGGC CTGTTCTGCT 660
GGATCAGTGA CCGAGCTTTC TGCGAGCTGC TGTATCTCTT CACTTCTCCC TACCTGCACT 720
GCATGTGCA CATCTCATC TGCCCTTCTG CCTACCTGGG CTGTGTATGC TTTGCTACT 780
TTGATGCTGC CTCAGAGATT CCTGAGCAAG GCCCTGTCTC CAGTTCTTGG CCCAATGAGA 840
AATGGGCTTT CATGTGTGTC CCTATGTGT CCTCTCTGTG TGCCAACAAG AAATCATCAG 900
TCAAGATCAC GTGATGGCAA GATGGTGGCT GGCCTCTCTG CTATCGCCCT CTCATGCAGT 960
GGGCTTCCTT TGCTAGGAAG ACAGCCAAAG GAGTTCGAAT AGTTGGGGTG TGGGCTATCT 1020
TTTCAAAAT CTATTGTCTG GGGCTCTTAA TTTCTTTAGT GTCTCTTSTA TGTAGGGATT 1080
TAAACTTTGT CATATGGTAC AAATATTCCC TGCCCCCTG CAGTTTCCCA TTTGTCTTTC 1140
AGTATGTTAA TATTTTGTG CCATCTGCTT TTTAACTTT CATGTGTGCA CATCTGTAAA 1200
TCTTTCTTT AGGATTTCTG GATTTTGTBT AATTTTTTAA AAGGTCCCTT CCTCTCTCCT 1260
AATGTGTCTG TGGACCACCT GGATTCCACT GTACAAGGGG AAAAGTGTCT ATTCTTTTCC 1320
CAAAGATGGA AATGGAGGG CTTAGGACA CTAGATGCAT CTCTCTCAGC ATCACTTCCA 1380
GATGCACTGA CTTGTGGGC TGCTCTCTTA ATGGCCATGG CAGAGCAGTC CCTTGGGGGA 1440
TCCAGCCCTG TACATGCAT CTCTTCTCTG AGAAAGCTGG CCTGCTCCAG ACCCCACCAT 1500
TCCAGGGGCG CCTTGGAGTG GACTCTACTG ATGACAGACA GACCCTCTGA GAGACAAGAC 1560
CCTCTGACTC TGTGATGGAA GATGCCAGAG ATTTTCTTTT GGGGTAAATG TCCTTAAACA 1620
AAACCAACA GATGAACAC ACACAGGACT TGTGGCTAAA AAGGCTAGTT TTCTACTTGC 1680
ATTTCTCAAC TAACCCAGG TTTACATGCA TCTGTGAATC CTTTACTAC TACCTCTGTG 1740
GAGAGATGGA GAGACTTCAG ATAAAGCTGA AGCTAATGAG TAAACCCCTC TCTGCCAAAA 1800
CCTACACTCC ACTTAGGCC CTCTTGAAG ATGAGCACAA TTTTAAATA CTGAGCACAA 1860
TTTTTAAATA CTGACATCAC TTCTCTTCTC CCCTCCACCC CAGCTCAGC AGCCTCAAA 1920
CTACAGAGAA GAAGAATTAT GGCATGAACA TTCCACAGA CCCACCATCT TTAAGACTTG 1980
ACCTCTGTAA GTTTACCAAA GGGCTCTCA CAATTGTGGT GGGGGTCTG GTTCAAAATT 2040

```

5 TGGAGCAAAAC ATGAAGTTT TGGAAACGTT TTCTCATTG AAGCCTCCAG TATGCTGTAC 2100
 TATTCTGGAA ATTACCTTCA AGAGTCTCAC TTCTTGTTC TGTGTGTTT TCTGTGGGCA 2160
 TCATGTTCTT CACCGTTGCA GTAGAAGGTG CTTTCTCGGT TTCCAGAGT ATCCACGGC 2220
 TCACCTTTCT CAAGTGTGG CAGTAGCTAT GCACCTACGG GCTGTTTGG GTGCTGGTG 2280
 CAGCAGGCGA AATCTGTGTC CTTCTGAATT TTCTCACCT AATGTGACAC TGGCTACAAT 2340
 GAATCTCTC TTCATCGGGC TGAATGAAG ATTCAGAAGC CATCTTCAAG GTGCATGGTG 2400
 GGAATTATCA ACCTCAGGGA TACTCATTT AACTCAGGCG TGTCCTGCTT TGTACATTG 2460
 CATGTGTGGG AGAGGGCAGG ACAGGTGTGT TCTCTGTGG GCAGGAGTCA TGTCACTGTC 2520
 10 CTACATATGT AAGAGTTGGG AAGGTGACGA TTCTTGACAC ATCCAGGAAAC TCTTACTCTA 2580
 GTTGAATTT GTACCAGATC CAGGTGAAA ACCCCAATAA GCAACTGAAT TTAGAGTTA 2640
 AAAATGAATG ACTTATGCT ACATCTGTG TTATCAAAT ATATAGGTG TTGAGAAGCA 2700
 GAACGCTGTT TGTAGTAAGA AATCTTGTG GAACCCAGT GTGTGAAGTA AATTGTATG 2760
 TATTAAATTT ATTTAAGGTT AAATTTATG CATTTACTTA ATAATATATG AGGTGGTGAA 2820
 AATGCAAAAT AACAAATGG TAATTTCCAA GGTAGAAAAA TTAGGTGTG AATGAATGTA 2880
 15 TATGTTGGT TATTATATAT ATAGAGAGAG TGTTTGTGT TTGAGTCAGG 2940
 GTCTGTCTCT ATTAACCCAG CTGGAGTGCA GTGGTGCACT CATGACTCAC TGTAGCTCT 3000
 GTCTCCAGG CTCAGGTGAT CCTCTCACCT CAGCCTCCCC AGGAGCTGGG AGTACTGGTG 3060
 CGTGCTTCCA CTCCAGCTCA ATTTTGTAT TTTTTCATAG AGATGGGGT TCACCATTTT 3120
 GCGCCAGGCT GTCTCAAAT TCTGGGCTCA AGAGATTGGC CCGCCTCGGC CTCCCAAGT 3180
 20 GCTGGGAATA GGCATGAGCC CTCACGCCG GCTCAAAAAA TATTTTAAAT ATGATCTTTG 3240
 AATTAAATAT TGTAGAAAT TCTAATGTAT CTCCTTGAGA CTTAGGAGGT TGATGGAACA 3300
 GAACCTGCTG TAAGTCTTCT GGGTTTCAAG TCTAGAATT TTTAAAGGCA AATATCAGCT 3360
 CATTCTTATT TTAGATGAG CTATCAGGC ATGGATTCTG GTCTCATCTA CTTTATGGTA 3420
 TAAATGCTCC AAGGTAGGG GTTTGGTATA TATTTTAAAG CCGGCTTTT TTTTITTTT 3480
 25 TTTTITTTT TTTTATGTT GAGAAGCAGA ATGTGCTTCT AGAACTGGT TTTAAAGAGA 3540
 TGAAGTGAAGA AAGAAATGTG GAATGGAGTA TATTGAGGA GGACAAACA TAACCTCACT 3600
 TTTGAACAGA AATCACTCTA GCTTGCCAGC ATGGGATGTA AACCAAGAGA GTAGAAATAT 3660
 ACCCATCTTA TTTTAAAGTT GGTTTATGGC ATCGCTCATA TATGTAAAAG CACTACAAAC 3720
 TCTTAAAGA AATTTGGGAA ACTACAGAGA AGTCAAGAAA AAAAAAAGT AACCCATATT 3780
 30 TCTATTGCC AGGTAAATC CTGGTTAATA TTTTGGTTG GTCTCTCTT TTTTCCCCC 3840
 AATATAGTTG TAAATAAATG ATGTCTTCA GAGTTGACAT TTATCTCTGA GCTTGAATGG 3900
 CATGTAAATG CCAAGTGTAT ATTTTTCAT GAAGTGTAGG TTTGGAATAC ACTAGAGTTA 3960
 GCTATATGCT TGAATGCTGA TCACCTGGAT CTGAGACTGA CTACTGAGTC TACCTTTTAA 4020
 ATCAAGCCTA ACATGAATGG GCTCCAAAAA GTAATGAATG TAATGTACT TTTTATGTTG 4080
 35 CCTCTGCACT TGGCTTGGTG AGTCATCATA AATAGCTGTT AAATATGTGA CTTTACAGAT 4140
 TTTGATATGT TCAGATTGTA AAAAATGAAT AGTTTATTTT ATTAATATGAT GGGCAGTCAA 4200
 GAATCTCCCT CC

Seq ID NO: 152 Protein sequence
 Protein Accession #: Bos Sequence

40 1 11 21 31 41 51
 | | | | |
 45 MGAPHWMDQL QAGSSEVDNC EDNYTIHPAI AEFYNTISNV LFFILPPICM CLPRQYATCF 60
 NSGIYLWL LVVVIGISVY FHATLSFLGQ MLDELAVLWV LMCALAMWFP RRYLPKIFRN 120
 DEGRKFKVVS VLSAVTTCLE FVKPAINNIS LMTLGVPCIA LLIAELKRCO NMRVFKLGLF 180
 SGLNWLALF CNISDRAFCE LLSFNFEPYL HOMHILICL AAYLGVCVFA YFDAASEIPE 240
 QGPVIRFNPEN EKNAFIGVPY VSLLCANKKS SVKIT

Seq ID NO: 153 DNA sequence
 Nucleic Acid Accession #: NM_001432.1
 Coding sequence: 167..676

55 1 11 21 31 41 51
 | | | | |
 60 TCACCTGCTT GATATTTCCT GTGTCAGAGG GACACAGCCA ACGTGGGGTC CCTTCTAGGC 60
 TGACAGCCGC TCTCCAGCCA CTGCGCGGAG CCGGCTGCTC CCGGCCCTGC CCGTGCCTC 120
 TCGCAGCGCG CCGTCCGCGA AGCCCTCAGG CCGGCTCCCA TCGCCGATGA CCGCGGGAG 180
 65 GAGGATGGAG ATGCTCTGTG CCGGCGAGGT CCGTCTGCTG CTGCTCTGCC TGGGTTTCCA 240
 TCTTCTACAG GCAGTCTCTA GTACCACTGT GATTCCATCA TGTATCCAG GAGAGTCCAG 300
 TGATAACTGC ACAGCTTTAG TTCAGACAGA AGACAATCCA CGTGTGGCTC AAGTGTCAAT 360
 AACAAAGTGT AGCTCTGACA TGAATGGCTA TTGTTTGCAT GGACAGTGCA TCTATCTGGT 420
 GGACATGAGT CAAACTACT GCAGGTGTGA AGTGGGTAT ACTGGTGTCC GATGTGAACA 480
 70 CTCTTTTTTA ACCGTCCACC AACCTTTAAG CAAAGAGTAT GTGGCTTTGA CCGTGAATCT 540
 TATTATTTTG TTTCTTATCA CAGTCTCGG TTCCACATAT TATTCTGCA GATGTTACAG 600
 AATCGAAAA AGTAAGAAC CAAAGAAGGA ATATGAGAGA GTTACCTCAG GGGATCCAGA 660
 GTTGGCGCAA GTCCTGAATG CCGCATCAA CTTATGGGCA GGGATAACAG TGTGCTTGGT 720
 TAATATTAAT ATTCATTTT ATTAATAATA TTTATGTTGG GTCAAGTGT AGGTCAATAA 780
 75 CACTGTATTT TANTGTACT GAAAAATGTT TTTATTTTG TTTTATTTT GACAGACTAT 840
 TTGCTAATGT ATAATGTGCA GAAAAATTTT AATATCAAAA GAAATTTGAT ATTTTATAC 900
 AAGTAATTT CTGAGCTAAA TGCTTCATTG AAGCTTCAA AGTTTATATG CCGGTGTCAC 960
 AGTCTTAGA AGTAAGCAAT TCCAGGTCA TAGCTCAAGA ATTGTAGCA AATGACAGAT 1020
 TTCTGTAGC CTATATATAT AGTCAAATCG ATTTAGTAAG TATGTTTTT AXTTCTCTCA 1080
 AATCAGTAT AATTGTTTG ACTGTACCAT GGTGTGATAT GTAGTTGGCA CCAATGGTATC 1140
 80 ATATATTAAC AATTAATGCA AATTAGAATT TGGGAGAAGC AAATATAGGT CCGTGTGTTAA 1200
 ACCTACACA TTTGAACAAA GCTAACCTTG GGGAGTCTAT GGTCTCTTCA CTCAGGTCTC 1260
 AGCTAATAT CAGTATATG AGGGGCACTG GACAGTTCCT TATGCCAACT CAGACTCTCT 1320
 ACAGGTACTA GTCACCTATC TACCAGATTG TGCCATATGA AATYGAATTG AAAAACAATT 1380
 TTCTGTAATC TTTTATTTAA GTAGTGGGCA TTTTATAGCT TCACATATGT CCTTTTGTGT 1440
 ATATTACAAC ATTTATGTGA GGTAAATATT GCTCAACAGA CAATTAGAAA AAGTCCACA 1500
 CTGAGGCTT AATTTTGTGC TTTTAAAGAA TATTTTAGA CTATTTCTT TTATAGGGGC 1560
 TTTGCTGAAT TCTAACATTA AATCACAGCC CAAATTTTGA TGGACTAATT ATTAATTTAA 1620
 AATATATGAA GACAATAATT CTACATGTTG TCTTAAGATG GAAATACAGT TATTTCTCT 1680
 TTTATTTCAAG GAAGTTTTAA CTTTAATACA GCTCAGTAAA TGGCTTCTTC TAGAATGTAA 1740

5	AGTTATGTAT	TTAAAGTTGT	ATCTTGACAC	AGGAAATGGG	AAAAAAGCTT	AAAAAAGCTT	AAAAAAGCTT	1800
	TGGTGTATTT	TTCCAAATGA	AAAATCTCAA	TTGAAAGCTT	TTAAATGTGA	TTAAATGTGA	TTAAATGTGA	1860
	CACACCTTCC	TGTGGAGGCT	GAGATGAAAA	CTAGGCTCA	TTTCTCTGAC	TTTCTCTGAC	TTTCTCTGAC	1920
	TTTTTGAAG	AGACAAAGAT	TTCTTCTGCA	CTCTGAGCCC	ATAGGTCTCA	ATAGGTCTCA	ATAGGTCTCA	1980
	AGGAGTATTT	TTGGGCTATT	GCATAAGGAG	CCACTGCTGC	CACCACTTTT	GGATTTTATG	GGATTTTATG	2040
	GGAGGCTCCT	TCATCGAATG	CTAAACCTTT	GAGTAGAGTC	TCCCTGGATC	ACATACCAGG	ACATACCAGG	2100
	TCAGGGAGGA	TCTGTTCTTC	CTCTACGTTT	ATCCTGGCAT	GTGCTAGGCT	AAACGAAAGC	AAACGAAAGC	2160
	ATAATAAGCC	ATGGCTGACC	TCTGGAGCAC	CAGGTGCCAG	GACTTGTCTC	CATGTGTATC	CATGTGTATC	2220
	CATGCATTAT	ATACCCCTGGT	GCAATCACAC	GACTGTCTATC	TAAAGTCTTG	GCCCTGGCCC	GCCCTGGCCC	2280
10	TTACTATTAG	GAAATAAACC	AGACAAAAAC	AAGTAAATAT	ATATGGTCTT	ATACATATTG	ATACATATTG	2340
	TATATATATT	CATATACAAA	CATGTATGTA	TACATGACCT	TAATGGATCA	TAGAATTGCA	TAGAATTGCA	2400
	GTCAATTGGT	GCTCTGCTAA	CCATTATAT	AAAACCTTAA	AACAAGAGAA	AAGAAAAATC	AAGAAAAATC	2460
	AATTAGATCT	AAACAGTTAT	TTCTGTTTCC	TATTTAATAT	AGCTGAAGTC	AAAATATGTA	AAAATATGTA	2520
15	AGAACACATT	TAAATATCTC	TACTTACAGT	TGGCCCTCTG	TGGTATGTTT	CACATCTGTT	CACATCTGTT	2580
	GATTCAACCA	ACCAAGGAGC	GAAAAATGCTT	AAAAAATGCTT	ACACAAACAA	CAAAAAATAC	CAAAAAATAC	2640
	ATATATACAA	CTATTACTTT	TTTTTTTTTT	CTTTTGTAGA	TGGAGTCTCG	CTCTGTTGCC	CTCTGTTGCC	2700
	CAGGTTGGGT	TGCTATGGCA	CBATCTCGGC	TCACTGCAAC	CTCACCCTCC	GGGTTCAAGA	GGGTTCAAGA	2760
	GATCCTCCTG	CCTCAGCCTC	CTGAGCAGCT	GGGACTACAG	GCGCATGCCA	CCATGCCCCG	CCATGCCCCG	2820
20	CTAATTTTGT	TATTTTATGT	AGAGGCGGCG	TTTCACCATG	TGGGCTAGGA	TGGTCTCAAT	TGGTCTCAAT	2880
	CTCCTAACCT	TGAGATCCAC	CCTCCACAGC	CTCCCAAACT	GCTGGGATTA	CAGGCGTGAG	CAGGCGTGAG	2940
	CCACCGCAGC	TAGCATTTAC	ATTAGGTATT	ACAAGTAATG	TAAAGATGAT	TAAAGTATAC	TAAAGTATAC	3000
	AGGAGGATCT	GATATAGTTA	TATGCAAGCA	CTATGCCCTT	TTATATTAAGT	GACTTGACAA	GACTTGACAA	3060
	TCGTGTCCCG	ATTTTAGTAT	GTGCAGGCGG	GCGATCTGGG	AATCAGTCCC	CTGTGGATAC	CTGTGGATAC	3120
25	CAAGTACAAA	CTGATTTTAT	TAACGCTTAC	TAGATGTGAG	GAGAGTCTGA	ATATTTTTCG	ATATTTTTCG	3180
	TGATCTTGGC	TGTTTCAAAA	AAATCTATTG	ACTTTTCAAT	AAATCAGCTG	CAATCCATTT	CAATCCATTT	3240
	ATTTCAATTA	CAAAAGATTT	ATTGTAAGCC	TCTCAATCTT	GGTTTTTTCG	TGATCTTTCG	TGATCTTTCG	3300
	GCAATGTCAAT	TCATAAAAGC	AAGTCATTTT	TGATTTTTC	ATCTTTAAGA	ATGCTTAAAA	ATGCTTAAAA	3360
	AAGCTAATCC	CTAAATATAGT	TAGATCTTTG	TAAATGCATA	TTAAATAATA	AAGTATGACC	AAGTATGACC	3420
30	CACATTAATT	TTTATGGGTC	AAAATAAGAC	AAAAATAATA	GTTTTAGTGA	GGATGGTGCT	GGATGGTGCT	3480
	GAGTAAACAT	AAAAACTGAT	TGCTCTCAG	CTGATGTGTC	CTGTACACAG	TGGGAAGATT	TGGGAAGATT	3540
	TTAGTTTCAA	CTTAGTCTAA	CTCCGCCATT	TTACAGATTT	CTCATTATAT	ATATTTCTAG	ATATTTCTAG	3600
	AAGGGGCTAT	GCAATATCAA	TGATTTGAGA	ACCAAGGCAA	CCACAAAGTC	ATAAATGCAT	ATAAATGCAT	3660
	AATTTATGGT	CTTCAACCAA	GGCCACATAA	TAAACCAATT	AATTTACTCT	TAAACCAAGG	TAAACCAAGG	3720
35	ATATTAAGTT	CTATACATAG	TACTCAAGGT	TAAACCTTAA	AATTAAGATT	TCCCTAACCT	TCCCTAACCT	3780
	TAACTTAAA	ATTGATATTA	TATTAACAT	ACATAATACA	ATGTAACCTC	ACTGTTCTCC	ACTGTTCTCC	3840
	TGAATATTTT	TGCTCTTAAT	CTCTCTGCCG	AAAGTCAAAG	TGATGGGAGA	ATTGGTATAC	ATTGGTATAC	3900
	TGGTATGACT	ACGCTCTAAG	TCAGATTTIT	ATTATGAGT	CTTTGAGACT	AAATTCATCT	AAATTCATCT	3960
	ACCACCAAGT	ATCAAAATCAA	CTTTTATGCA	GCAATATAT	GATTCATAGT	TCTGACTTTT	TCTGACTTTT	4020
40	GTAAATTTCA	GTAATGCACT	TTTTAAAAAC	CTGTATCTGA	CCACCTTTGT	AAATTTTGTCT	AAATTTTGTCT	4080
	CCAATATCCA	TCTGTAGATC	TTTTGAAAAA	AAAGTTTFTA	ATTGTATGCC	CAATATATTC	CAATATATTC	4140
	TGACCGTTAA	AAATTTCTGT	TTCAATATGG	AGAGGGGGGA	GTAATGACTT	GTACAAACAG	GTACAAACAG	4200
	TATTTCTGGT	GTAATTTTGA	ATGTTTTTAA	AAAGAGTAA	TTCAATTTAA	TATCTGTAT	TATCTGTAT	4260
	TCAAAATTGA	TGATGTTAAA	TGTAATATAA	TGATTTTCT	TTTTATTTTG	CATCTCTGTA	CATCTCTGTA	4320
45	TGACCTTTT	TAGTTTGA	GAGCCATTTT	GGTAAACGGT	TTTTATTTAA	GATGCTATGG	GATGCTATGG	4380
	AACATAAAGT	TGTATTGCAT	GCAATTTAAA	GTAACCTATT	TGACTATGAA	TATTATCGGA	TATTATCGGA	4440
	TTACTGAATT	GTATCAATTT	GTTTGTGTTC	AAATACAGCT	TTGATTAATG	TGTACCTTAA	TGTACCTTAA	4500
	GATATTGAAG	GAGAAATAG	ATAATTTACA	AGATATTATT	AAATTTTATT	TATTTTCTCT	TATTTTCTCT	4560
50	GGGAATTGAA	AAAAATTGAA	ATAAATAAAA	ATGCATTGAA	CATCTTGATC	TCAAAATCTT	TCAAAATCTT	4620
	CACGTGAC							

Seq ID NO: 154 Protein sequence
Protein Accession #: NP_001423.1

55	1	11	21	31	41	51	
	MTAGRMEML	CAGRPALIL	CLGPHLLQAV	LSTTVIPSCI	PGESSINCTA	LVQTEDNPRV	60
	AQVSIITKCS	DMNGVCLHQ	CYLVDMDSQN	YCRCEVGYTG	VRCEHFLIV	HQPLSKSEYVA	120
	LTVILLILPL	ITVVGSTYYF	CRWYRNRKSK	EPKKEYERV	SGDFELPQV		

Seq ID NO: 155 DNA sequence
Nucleic Acid Accession #: NM_013282.2
Coding sequence: 85..2466

65	1	11	21	31	41	51	
	CGACTCCTTA	GAGCATGGCA	TGGCTCAGAG	GTGCTGGTAA	AACTGATGGG	GGTTTTTGCT	60
	GTCCCTCCCC	TCAGCGCCGA	CACCATGTGG	ATCCAGGTTT	GGAACATGGA	CGGGAGGCAG	120
	ACCCACACGG	TGAGCTCGCT	GTCCAGGCTG	ACCAAGGTGG	AGGAGCTGAG	GCGGAAGATC	180
70	CAGGAGCTGT	TCCACGTGGA	GCCAGGCTGT	CAGAGGCTGT	TCTACAGGGG	CAACAGATG	240
	GAGGACGGCC	ATACCTCTCT	CGACTACGAG	GTCCGCTCTG	ATGACACCAT	CCAGCTCTCT	300
	GTCCGCCAGA	GCTCTGTGCT	CCCCACAGC	ACCAAGGAGC	GGGACTCCGA	GCTCTCCGAC	360
	ACCGACTCCG	GCTCTGTGCT	GGGCCAGAGT	GAGTCAGACA	AGTCTCTCC	CCACGGCGAG	420
	CGCGCCCGCG	AGACTGACAG	CAGGCCAGCC	GATGAGGACA	TGTGGGATGA	GACGGAAITG	480
75	GGGCTGTACA	GCTCTCAATG	GTAAGTCTGT	GCTCGGAGCA	CGAACATGGA	GCCGTGGTTT	540
	GAGGCGCAGG	TGCTCAGGGT	GACGCGGAGG	GCCCCCTCCC	GGGACGAGCC	CTGCAGCTCC	600
	ACGTCCAGGC	CGGCGCTGGA	GGAGGACGTC	ATTTACCAAG	TGAAATACGA	CGACTACCGG	660
	GAGAACCGCG	TGCTCCAGAT	GAACTCCAGG	GAGTCCGAGG	CGCGCGCCCG	CACCATCATC	720
	AAGTGGCAGG	ACCTGGAGGT	GGGCCAGGTG	GTCATGCTCA	ACTACACCCC	CGACAAACCC	780
80	AAGGAGCGGG	GCTCTGTGTA	CGACGCGGAG	ATCTCCAGGA	AGCGCGAGAG	CAGGACGGCG	840
	CGGGAACCTC	ACGCCAACGT	GGTGTGGGGG	GATGATTTCT	TGAACGACTG	TCCGATCATC	900
	TTCTGTGGAG	AGGTCTTCAA	GATTGAGCGG	CCGGGTGAAG	GGAGCCCCAT	GGTTGACATC	960
	CCCATGAGAC	GGGAGAGCGG	GCGGCTCTGC	AAGCACTGCA	AGGACGACGT	GAACAGACTC	1020
	TGCGGGGTCT	GCGGCTGCGA	CCTGTGCGGG	GGCCGCGCAG	ACCCGACCAA	GCAGCTCATG	1080
	TGCGATGAGT	GCGACATGGC	CTTCCACATC	TACTGCTCTG	ACCCGCCCCC	CAGCAGTGTT	1140

5
 10
 15
 20
 25
 30
 35
 40
 45

```

CCCAGCGAGG ACAGTGGTGA CTGCCCTGAG TGCCGGAATG ATGCCAGCGA GGTGGTACTG 1200
GCGGAGAGAG GGTGAGAGAG GAGCAAGAAG AAGGCGAAGA TGGCCTCGGC CACATCGTCC 1260
TCACAGCGGG ACTGGGCGAA GGGCATGGCC TGTGTGGGCC GCACCAAGGA ATGTACCATC 1320
GTCCCGTCCA ACCACTACGG ACCCATCCCG GGGATCCCGG TGGGCACCAT GTGGCGGTTT 1380
CGAGTCCAGG TCAGCGBAGT GGGTGTCCAT CGGCCCCACG TGGCTGGCAT ACACGGCCCG 1440
AGCAACGAGG GAGCGTACTC CTTAGTCTTG GCGGGGGGCT ATGAGGATGA CBTGGACCAT 1500
GGGAATTTT TCACATACAC GGTAGTGGT GGTGAGATC TTTCCGCAAA CAAGAGGACC 1560
GCGGAACAGT CTTGTGATCA GAAACTCACC AACACCAACA GGGCGCTGGC TCTCAACTGC 1620
TTTGTCCCA TCAATGACCA AGAAGGSCCC GAGGCCAAGG ACTGGCGGTC GGGGAAGCCG 1680
GTCAGGGTGG TGCCTCAATG CAAGGGTGGC AAGAATAGCA AGTACGCCCC CGCTGAGGGC 1740
AACCCTACG ATGGCATCTA CAAGGTGTG AAATACTGGC CGGAGAGGGG GAAGTCCCGG 1800
TTCTCGTGT GGCCTACCT TCTGCGGAGG GACGATGATG AGCCTGGCCC TTGGAAGGAG 1860
GAGGGGAGG ACCGGATCAA GAAGCTGGGG CTGACCATGC AGTATCCAGA AGGCTACCTG 1920
GAAGCCCTGG CCARCCGAGA GCGAGAGGAG GAGAACAGCA AGAGGGAGGA GGAGGAGCAG 1980
CAGGAGGGGG GCTTCGCGTC CCCGAGGAGG GGCAAGGGCA AGTGAAGCG GAAGTCCGCA 2040
GGAGGTGGCC CGAGCAGGCG CGGTCCCGG CGCCGGACAT CCAAGAAAC CAAGGTGGAG 2100
CCCTACAGT TCACGCGGCA GCAGAGCAGC CTCATCAGAG AGGACAAGAG CAACGCCAAG 2160
CTGTGGATG AGGCTCGGC GTCACTCAAG GACCGGCGCG CGAGCGCGAG CCGTTCGAG 2220
TTGTTCCTGA GTAAAGTGA GGAGACCTTC CAGTGTATCT GCTGTCAAG GCTGGTGTTC 2280
CGGCCCATCA CGACCGTGTG CCAGCAACAC GTGTGCAAGG ACTGCTTGA CAGATCTTT 2340
CGGGCAGAGG TGTTCAGCTG CCCGCTCTGC CGCTACGACC TGGGCGCGAG CTATGCCATG 2400
CAGGTGAACC AGCCTCTGCA GACCTCTCTC AACCACTCT TCCCGGCTA CGGCAATGGC 2460
CGGTGATCTC CAAGCAGTTC TCGACAGGCG TTTGTCTGAA AAGTGTCTG AGGCTCTGTT 2520
CATCGGCACT GATTTGTGTT TTAGTGGGCT TAACITAAAC AGGTAGTGT TCTCTCGTTC 2580
CCTAAAAGG TTTGTCTTCC TTTTITTTTA TTTTATTTT TCAAACTAT ACATTTTCAG 2640
GAATTATGT ATCTGGGCTA AAAGTTGGAC TTCTCAGTAT TGTGTTTGT TCTTTGAAA 2700
CATAAAAGCC TCAGATTTCT CGACAAAACA ACACAGATT TTTTAAAGT GGAATCAGAA 2760
ACTACGTGTT GTGAGGGCTG TTGATGTTTC TGGTGTCAAG TTCTCAGAA TTGCTGCCAC 2820
CAACTCTTA GAAGGCGGAC AGGATCAGTC CTCTCTAGG GTCTGCGCC CCAAGGTGAG 2880
AGCAAGCATC TTCTGACAG CATTTTGTCA TCTAAAGTCC AGTGACATGG TTCCCGTGG 2940
TGGCCCGTGG CAGCCCGTGG CATGGGTGG CTCAGCTGTC TGTGAAAGT GTTGCAAGGA 3000
AAAGAGGAAA CATCTCGGCG CTAGTTCAAA CCCTTGCCCT AAAGCTATCC CCGACAGAG 3060
TGCTTAGCTT CAGAGATCCG CGTGAAAGT CCTCTGCCCA CGAGAGCAGG GAGTTGGGGC 3120
CAGCAGAAA TGCCCTCAAG GGGACTCTGC TCCAGTGGG GCCAGGCGTG TGACTGACG 3180
TGTCCGACGA AGCGGGCCAC GAGCGGACGC CAGCACAGCA AGTCACGTGC AAGTGCCTTT 3240
GATTCGTTCC TTCTTTCTAA AGACGACAGT CTTTGTGTT AGCACTGAAT TATGAAAT 3300
GTCAACAGAG TTCTAGAAAC TGCGGTCAAT CAGTCTCTCC TGACACCGGA TGGGTGCTTG 3360
GGAAACGTTT GAGCCTTATA GATCATTAC ATTCAATTT TTTAACTCAG CAAGTAGAAA 3420
CTTACAGAG GGTTTTTTTT TAATTTTTTT TTCTCTTAAT GAACACATT TCTAAATGAA 3480
TTTTTTTTGT AGTTACTGTA TATGTACCAA GAAAGTATA ACGTTAGGAT TTGGTTGTTT 3540
TTGTTTTTGT ATTTTTTTTC TTTTGAAAGG GTTTGTTAAT TTTTCTAAT TTACCAAGT 3600
TTGCAGCCTA TACCTCAATA AAACAGGGAT ATTTTAAATC ACATACCTGC AGACAAACT 3660
GAGCAATGTT ATTTTAAAG GGTTTTTTTC AACTCTTAT TCTTAGATTA TTAATGTATT 3720
AGGGAAGAT GAGACAAATT TGTGTAGGCT TTTTCTAAG TCCAGTACTT TGTCAGATT 3780
TTGATTTCTC AGAATAAATG TTTTTCACAG ATTGAAAAA AAAAAAAA
  
```

Seq ID NO: 156 Protein sequence
 Protein Accession #: NP_037414.2

50
 55
 60
 65

```

1 11 21 31 41 51
MWIQVRTMG RQTEVDLSL RLTKVEELRR KIQLFHEVEP GLQLFHYRQK QMEDGETLFD 60
YVRLNDTIQ LLVRQSLVLP HSTKERDSEL SDTDSGCCLG QSESDKSSTH GEAAAEEDSR 120
PAEDMMDET ELGLYKVNEY VDARDTMMGA WFEAQVVRVT RKAPSRDEPC SSTGREALEE 180
DVIYEVIYDD YPENGUVQMN SRDVRARART IIKNQDLEVG QVVMILWYNDP NPKBRGFWYD 240
AETSRKRETR TARELYANVU LGDDSLNDCR IIFVDEVFKI ERFGGSGPMV DNFMRKSGP 300
SKHKCKDVN RLRCVACACH CGGRQDPDKQ LMCDBCMAY HIYCLDPLLS SVPSDEBWC 360
PECHNDASEV VLAGERLRES KICKAKMASAT SSSQRDWGKG MACVGRTEKC TIVPSNHVYP 420
IPGIPVGTMW RFRVGVSESG VHRPHVAGIH GRSDGAYSL VLAGGYEDDV DEGNFTTGT 480
SGGRDLSENK RTASQSCDQK LINTNRALAL NCFAPIMDQE GAEAKDWRSG KPVVRVVRNVK 540
GGRNSKYAPA EGNRYDGIYK VVKYNPERGK SGFLVWRYLL RRDDDEPGPW TKEGRDRIKK 600
LGLIMQYPEG YLEALANRRR EKENSKEEES BQEGGFASP RTGKGKWKRE SAGGSPBRAG 660
SPRRTSKKTK VEFYSLTAQQ SSLIREDKSN AKLWNEVLAS LKDRPASGSP FQLFLSKVES 720
TFQCICQQL VFRPIITVQ HNVCKDCILDR SFRAQVFSCP ACRYDLGRSY AMQVWQPLQT
  
```

Seq ID NO: 157 DNA sequence
 Nucleic Acid Accession #: NM_000756.1
 Coding sequence: 186..776

70
 75
 80

```

1 11 21 31 41 51
AGAAACTCAG AGACCAAGTC CATTGAGAGA CTGAGGGGAA AGAGAGGAGA GAAAGAAAAA 60
GAGAGTGGGA ACAGTAAAGA GAAAGGAAGA CAACCTCCAG AGAAAGSCCC CGGAGACGTC 120
TCTCTGCAGA GAGGCGGCG CACCCGCTC ACCTGCGAAG CGCCTGGGAA GCGAGTGCCC 180
CTAACATGCG GCTGCCGCTG CTTGTGTCCG CGGAGTCTCT GCTGGTGGCT CTCTGCCCCT 240
GCCCGCATG CAGGGCGCTC CTGAGCGCG GCGCGGTCCC GGGAGCTCGG CAGGCGCGCG 300
AGCACCTCA GCGCTTGGAT TTCTTCCAGC GCGCGCGCA GTCCGAGCAG CCCCAGCAGC 360
CGCAGGCTCG GCGGCTCTG CTCCGCAATG GAGAGGAGTA CTCTCTCGC CTGGGGAACC 420
TCAACAGAG CCGGCGGCT CCGCTTTCG CCGCTCTCT GCTCTCTCG GAGGACAGCG 480
GCAGCGGCC TTGCGCGGAA CAGSCGACCG CCAACTTTTT CCGGTGTTC CTGACGACG 540
TGCTGCTGCT TCGGCGCTCG CTGACAGCC CCGCGGCTCT CGGAGAGCG GCGCTAGGA 600
ATGCCCTCG GCGCCACCG GAGGCACCG AGAGAGAAAG GCGGTCCGAG GAGCCTCCCA 660
TCTCCCTGGA TCTCACCTC CACTCTCTC GGAAGTCTT GGAATGGCC AGGGCCGAGC 720
AGTTAGCTA GCAAGCTCAC AGCAACAGGA AACTCATGGA GATTATTGG AATAAAAA 780
  
```

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

GTGCGTTTGG CCAAAAAGAA TCTGCATTGA GCACAAAAAA AATTTAAAAA AATACAGTAT 840
TCTGTACCAT AGCGCTGCTC TTATGCCATT TGTTTATTTT TATATAGCTT GAAACATAGA 900
GGGAGAGAGG GAGAGAGCCY ATACCCCTTA CTTAGCATGC ACAAGTGTGA TTCACGTGCA 960
GCAGCAACAC AATGTTATTG GTTTTGTCTA CTTTATGTTT CCGTTTCCAG GTGTTTATAG 1020
TGTGTGTTTA AAGAGAATGT AGACCTGTGA GAAAAAGTTT TGTGTAAGAA AGCAGACAGA 1080
AGTCACTCAA TGTGTTTGTG TGTGGTCTGA GCCAAAGAGA ATGCCATTCT CTTGGGTGGG 1140
TAAGACTAAA TCTGTAAGCT CTTTGAACA ACITTTCTCT GTAAACGTTT CAGTAATAAA 1200
ACATCTTTCC AGTCCCTGGT CAGTTTGGTT GTGTAAGAGA ATGTTGAATA CTTATATTTT 1260
TAATAAAAGT TGCAAAGGT

Seq ID NO: 158 Protein sequence
Protein Accession #: NP_000747.1

15
20
25
30
35
40
45
50
55
60
65
70
75
80

1 11 21 31 41 51
MRPLPLVSAG VLLVALLPCP PCRALLSRGP VEGARQAPQH PQPLDFDFPP PQSEQPQQPQ 60
AREVLLRMGE EYFLRLGLNL KSPAAPLSPA SSILAGSGSG RPSPEQATAN FFRVLLQQLL 120
LPRRLDSPRA ALAERGARGA LGGHEBAPER ERRSEEPFIS LDLTFHLLRE VLEMARABQL 180
AQAHSNRRL MBTIGR

Seq ID NO: 159 DNA sequence
Nucleic Acid Accession #: NM_001200.1
Coding sequence: 325..1514

1 11 21 31 41 51
GGGACTTCT TGAACCTGCA GGGAGAATAA CTTBGCACCC CCACCTTGGG CCGGTGCTT 60
TGCCCCAGCG GAGCCTGCTT CGCCATCTCC GAGCCCAACC GCGCCCTCCAC TCCTCGGCTT 120
TGCCCGACAC TGAGACGCTG TTCCACGCGT GAAAAAGAGAG ACTGCGCGGC CGGCACCCGG 180
GGAAGAGAGG AGGCAAGAA AAGGAACGGA CATTCGGTCC TTGCGCCAGG TCCTTTGACC 240
AGAGTTTTC CATGTGGAGC CTCITTCAT GGACGTGTCC CGCGTGTCTT CTTAGACGGA 300
CTGCGGTCTC TAAAGGTGCB ACCATGCTGB CCGGGACCCB CTGTCTTCTA GCGTGTCTGC 360
TTCCCGAGGT CCTCCTGGGC GCGCGGCTG GCTCTGTTC GAGCTGTGGC CGCAGGAAGT 420
TCCCGCGGC GTGCTCGGGC CGCCCTCAT CCCAGCCCTC TGACGAGGTC CTGAGCGAGT 480
TCGAGTTCG GCTGCTCAGC ATGTTGCGCC TGAACAGAG ACCCACCCCC AGCAGGAGC 540
CGGTGCTGCC CCCCTACATG CTAGACCTGT ATCGCAGGCA CTCAGGTGAG CCGGGCTCAC 600
CCGCCCCAGA CCACCGGTTG GAGAGGCGAG CCAGCCGAGC CAACACTGTG CGCAGCTTCC 660
ACCATGAAGA ATCTTTGGAA GAATACAGG AAACGAGTGG GAAACAACCC CGGAGATTCT 720
TCTTTAATT AAGTTCTATC CCCACGAGG AGTTTATCAC CTCAGCAGAG CTTAGGTTT 780
TCCGAGACA GATGCAAGAT GCTTTAGGAA ACAATAGCAB TTCCATCAC CGAATTAATA 840
TTTATGAAT CATAAAACCT GCAACAGCCA ACTGGAATT CCCCCTGACC AGACTTTTG 900
ACACCGAGT GGTGAATCAG AATGCAAGCA GGTGGGAAG TTTTGATGTC ACCCCCGCTG 960
TGATGCGGTG GACTGCACAG GGCACGCCA ACCATGGATT CGTGGTGAA GTGCGCCACT 1020
TGAGGAGAGA ACAAGGTTTC TCCAGAGAC ATGTTAGGAT AAGCAGGCTT TTGACCAAG 1080
ATGAACACAG CTGCTCAGAG ATAAGGCCAT TGCTAGTAAC TTTTGGCCAT GATGGAAAG 1140
GGCATCCTCT CCACAAAAGA GAAAAACGTC AAGCCAAACA CHAACAGCGG AAACGCTTGA 1200
AGTCCAGCTG TAAGAGACAC CCTTGTGACG TGGACTTCAG TGACGTGGGG TGGAATGACT 1260
GGATTGTGGC TCCCCCGGGG TATCAGCCTT TTTACTGCCA CGGAGAAATG CCTTTTCTC 1320
TGGCTGATCA TCTGAATGCC ACTAATCATG CCATTGTTCA GACGTGGTTC AACTCTGTTA 1380
ACCTTAAGAT TCCTAAGACA TGCTGTGTCC CGACGAACT CAGTGTCTAT TCGATGCTGT 1440
ACCTTGACA GAATGAAAG GTTGATTAA AGAACTATCA GGACATGGTT GTGGAGGGTT 1500
GTGGGTGTG CTAGTACAGC AAAATTAAAT ACATAAATAT ATATATA

Seq ID NO: 160 Protein sequence
Protein Accession #: NP_001191.1

1 11 21 31 41 51
MVAGTRCLLA LLLPQVLLGG AAGLVEPLGR RKFAAASSGR PSSQPSDEVL SEFKILLISM 60
FGLKQRPPTS RDAVVPFVYL DLYRRHSGQP GSPAPDHRLE RAASRANTVR SFHHERSLEE 120
LPETSGKTR RFFFNLSIP TERFITSABL QVEREQWQDA LGNNSSPFRH INIYEIHKPA 180
TANSKEFVTR LLDT

Seq ID NO: 161 DNA sequence
Nucleic Acid Accession #: NM_001999.2
Coding sequence: 1..8736

1 11 21 31 41 51
ATGGGGAGAA GACGGAGGCT GTGTCTCCAG CTCTACTTCC TGTGGCTGGG CTGTGTGGTG 60
CTCTGCGGCG AGGGCAAGGC CGGCCAGCCT CAGCCTCTTC CGCCCAAGCC GCGCCGAGCC 120
CAGCCGCGCG CGCAACAGGT TCGGTCCGCT ACAGCAGGCT CTGAAGGCGG GTTCTAGCG 180
CCCGAGTATC CGAAGAGGGG TGCCGCGAGT GCCAGCCGCG TCCGCGCGCG AGGACAGCAG 240
GACGTGCTCC GAGGGCCCAA CGTGTGCGGC TCACAGATCC ACTCCTACTG CTGCGCTGGA 300
TGGAAGACGC TCCCTGAGAG AAACCAATGC ATTTGTCGGA TTTGTAGAAA TAGTTGTGGA 360
GATGATTTT GTTCCCGTCC TAACATGTGT ACTTGTTCCT GTGGGCAAA ATCATCAACC 420
TGTGATCAA AATCAATCA GCAGTGCAGT GTGAGATGCA TGAATGGTGG GACCTGTGCA 480
GATGACCACT GCCAGTGCCA GAAAGGATAT ATTGGAACCT ATTGTGGACA ACCGTCTGT 540
GAAATGGAT GTCAACAATG TGGACGTGCG ATGCCCAAC CGTGTGCTTG TGTATATGGG 600
TTCACTGTGC CACAGTGTGA AAGAGATTAC AGGACAGGCC CGTGTTCAC TCAGGTCAAC 660
AACCAATGAT GCCAAGGACA GCTGACAGGC ATTGTCTGCA CGAAGACTCT GTGCTGTGCG 720
ACCACTGTGC GGGCGTGGGG CCATCCCTGT GAGATGTGTC CAGCCAGGCC TCAGCCCTGC 780
CGAAGGGGTT TCATCCCAAA CATCCGCACT GGAGCTTGCC AAGATGTTGA TGAATGCCAG 840
GCTATCCAG GGAATATGCCA AGGAGGAAAC TGTATCAATA CAGTGGGCTC TTTTGAATGC 900

	AGATGCCCTG	CTGGTCACAA	ACAGAGTGAA	ACTACTCAGA	AATGTGAAGA	CATTGATGAG	960
	TGCAGCATCA	TTCTCTGGAT	ATGTGAAACT	GGTGAATGTT	CCAACACCGT	GGGAAGCTAT	1020
	TTTTGTGTTT	GTCCACGTGG	ATATGTAACC	TCAACAGATG	GCTCTCGATG	CATCGATCAG	1080
5	AGAACAGGCA	TGTGTTTCTC	GGGCTCTGGT	AATGGCCGCT	GTGCACAAGA	GCTCCCGGGG	1140
	AGAATGACGA	AAATGCACTG	CTGCTGTGAG	CCTGCGCGCT	GCTGGGGCAT	CGGAACCAAT	1200
	CCTGAAGCCT	GTCTGTCTAG	AGGTTCTGAG	GAATATGCGA	GACTTTTGAT	GGATGGACTT	1260
	CCAATGGGAG	GAATTCACAG	GAGTGTCTGT	TCCAGACCTG	GAGGCACTGG	GGGAATGGC	1320
	TTTGGCCCAA	GTGGCAATGG	CAATGGCTAT	GGCCACAGGAG	GGACAGGCTT	CATCCCAATC	1380
	CCTGGAGGCA	ATGGCTTTTC	TCCTGGCGTT	GGGGGAGCCG	GTGTGGGGGC	CGGGGGACAG	1440
10	GGACCTATCA	TCACCTGGACT	AACAATTCTG	AACCAGACAA	TAGATATCTG	TAAGCATCAT	1500
	GCTAACCTTT	GTTTAAATGG	ACGCTGTATA	CCAACCTGCT	CAAGCTACCG	ATGTGAATGC	1560
	AACATGGGTT	ATAAGCAGGA	TGCAAAATGGA	GATTGTATAG	ATGTTGATGA	ATGCACATCA	1620
	AATCCCTGCA	CTAATGGAGA	TTGTGTTAAC	ACACCTGGTT	CCTATTATTG	TAAATGTCAT	1680
	GCTGGATTCC	AGAGGACTCC	TACCAAGCAA	GCATGCAATG	ATATTGATGA	GTGCATCCAG	1740
15	AATGGGTTTC	TTTGTAAAAA	CGGTGATGCG	GTGAACTCAG	ATGGGAAGTTT	CCAGTGCATT	1800
	TGCAATGGCG	GCTTTGAATT	AACATCAGAT	GGAAAAAACT	GTGTTGATCA	TGATGAATGT	1860
	ACAACCTGCA	ACATGTGTTT	GAATGGAAATG	TGCATCAATG	AAGATGGCAG	CTTCAAGTGC	1920
	ATCTGCAAAC	CAGGATTGTG	CTTGGCTCCA	AATGGGCGTT	ACTGTACTGA	TGTGTATGAA	1980
20	TGCCAGACCC	CAGGAATCTG	CATGAATGGG	CACCTGCATCA	ACAGTGAAGG	GTCTTTCCGC	2040
	TGTGATGTGC	CCCCAGGCGT	GGCTGTGGGC	ATGGATGGAC	GTGTGTGTGT	TGATACTCAC	2100
	ATGGCAGTGA	CCGTCTATGG	AGGAATCAAG	AAAGGAGTGT	GTGTGCGTCC	TTTCCCGCGT	2160
	GCAGTGAAGA	ACTGCGGATT	CTGCTGTGCC	AATCCAGACT	ATGTTTGTGG	AGAACCTTGC	2220
	CAGCCATGCC	CTGCAAAAAA	TTCCAGCTGA	TTCCACGGCC	TTTGTAGTAG	TGGAGTAGGT	2280
25	ATCACTGTGG	ATGGAAGAGA	TATCAATGAA	TGTGCTTTGG	ATCCTGATAT	ATGTGCCAAT	2340
	GGGATTGTTC	TTTGTAAAAA	TGTTAGTTAC	CGTTGTAATT	GCAACAGTGG	CTATGAACCA	2400
	GATGCTCTTG	GAGGAAACTG	TATTGACATT	GATGAATGTT	TAGTAAACAG	ACTGCTTTGT	2460
	GATAAGGAGT	TGTAGCGAAA	CAGCCAGGGA	AGTTACAGCT	GTACGTGCCC	ACCAGGATAT	2520
	GTGTTCAAGA	CTGAGACAGA	GACCTGTGAA	GATATAAATG	AATGTGAAGG	CAACCCATGT	2580
30	GTCAATGGGG	CCTGCAGAAA	CAACCTTGGG	TCTTTCAATT	GTGAATGTTT	GCCCGGCAGC	2640
	AAACTCAAGT	CCACAGGATT	GATCTGTATT	GACAGCCTGA	AGGGGACCTG	TTGGCTCAAC	2700
	ATCCAGGACA	CGCGCTGTGA	GGTGAATATT	AATGGAGCCA	CTCTGAAATC	TGAATGCTGT	2760
	GCCACCCGTG	GAGCCGCTCT	GGGGAGCCCT	TGTGAGCGGT	GTGAACATGA	TACAGCTTGC	2820
	CCAAGAGGGC	TTCGCCAGAT	TAAAGGTGTT	ACBTGTGAAG	ATGTTAATGA	GTGTGAGGTG	2880
35	TTCCCTGGCG	TTTGTCCAAA	TGGAGCCTGT	GTCAACAGTA	AGGGATCTTT	TCATTGGGAG	2940
	TGCCCTGAAG	GCCTTACGTT	GGATGGGACT	GGCCGTGTAT	GTTTGGATAT	TCGCATGGAG	3000
	CAGTGTACTT	TGAAGTGGGA	TGAAGATGAA	TGCATCCACC	CCGTTCCTGG	AAAGTTCCGC	3060
	ATGGATGCCCT	GCTGCTGTGC	TGTGGGGGCH	GCTTGGGGCA	CGAGTGTGTA	GGAGTGCCCT	3120
	AAACCTGGCA	CCAAGGAATA	CGAGACACTG	TGCCCCGCCG	GGGCTGGCTT	TGCTAACCGA	3180
40	GGGATGTTTC	TTACTGGGCG	GCCATTTTAC	AAAGACATCA	ATGAATGCAA	AGCATTTCCT	3240
	GGGATGTGCA	CTTATGGGAA	GTGCAGAAAT	ACAATCGGAA	GCTTCAAAATG	CCGTTGCAAT	3300
	AGTGGCTTTG	CTCTAGACAT	GGAGGAAAGA	AACTGCACGG	ACATCGACGA	GTGCAGGATT	3360
	TCCTCTGAGC	TCTGTGGGCG	TGGAATCTGC	GTCAATACAC	CGGGCAGCTT	TGAGTGGGAG	3420
	TGCTTCGAAG	GCTATGAAGG	TGGCTTTCATG	ATGATGAAGA	ACTGCATGGA	CATTGACGGA	3480
45	TGTGAACGTA	ACCTCTCTCT	TTGTAGGGGT	GGCACCTGTG	TGAACACTGA	GGGCAGCTTT	3540
	CAGTGTGACT	GCCCACTGGG	ACACGAGCTG	TCACCATCCC	GTGAGGACTG	TGTGATATT	3600
	AATGAATGCT	CCCTGAGTGA	CAATCTCTGC	AGAAATGGAA	AATGTGTGAA	CATGATTGGA	3660
	ACCTATTCAG	GCTCTTGCAA	TCCTGGATAT	CAGGCTACGC	CAGACCGCCA	GGGCTGTACA	3720
	GATATTGATG	AATGTATGAT	AATGAACGGA	GGCTGTGACA	CCCACTGCAC	AAATTTCAGG	3780
50	GGAGGCTACG	AATGCAGCTG	CAGTGAGGGT	TATGCCCTGA	TGCCAGATGG	GAGATCTGTT	3840
	GCAGACATGG	ATGAATGTGA	AAACAATCCT	GATATCTGTG	ATGGCGGCCA	GTGTACCAAT	3900
	ATTCCTGGAG	AGTATCGCTG	CCCTCTGCTAT	GATGGCTTCA	TGGCTTCCAT	GGACATGAAA	3960
	ACATGCAATG	ATGTCAATGA	ATGTGACCTA	AATTCAAATA	TCTGCAATGT	TGGGAATGT	4020
	GAGAACACAA	AGGGATCCTT	CATTTGCCAC	TGTCACTGCG	GTTACTCAGT	GAAGAAGGGG	4080
55	ACCACAGGAT	GTACAGATGT	GGATGAGTGT	GAATTGGTGG	CTCATACACTG	GGACATGCTT	4140
	GGCTCATGTC	TGAATATCCC	AGGAAGCTTC	AAGTGTAGCT	GCAGAGAAGG	CTGGATTGGA	4200
	AACGGCATCA	AGTGTATGGA	TCTGGAGGAA	TGTTCTAATG	GAACCCACCA	GTGTAGCATC	4260
	AATGCTCAGT	GTGTAAATAC	CCCGGCTCA	TACCGCTGTC	CCTGCTCCGA	AGGTTTCACT	4320
	GGTATGCGCT	TTACCTGCTC	AGATGTTGAT	GAGTGTGCMG	AAACATAAA	CCTCTGTGAG	4380
60	AACGGACAGT	GCTTTAATGT	CCCGGCTGCA	TATCGCTGCG	AGTGTGAGAT	GGGCTTCACT	4440
	CCAGCCTCAG	ACAGCAGATC	CTGCCAAGAT	ATTGATGAAT	GCTCCTTCCA	AAACATTGTT	4500
	GTCTCTGGAA	CATGTAATAA	CCTGCTCTGA	ATGTTTCATT	GCATCTGCGA	TGATGGTTAT	4560
	GAATTGGACA	GAACAGGAGG	GAACCTGTACA	GATATTGATG	AGTGTGCGAG	TCCTATAAAC	4620
	TGTGTCAATG	GCTTATGTGT	CAACACGCCCT	GCTGCTCATG	AGTGTACATG	CCACCCGAT	4680
65	TTTCAGTTGA	AOCCAACTGG	TGTGGGTTGT	GTGACAAAC	GTGTGGGCAA	CTGCTACCTG	4740
	AAQTTTGGAG	CTCGAGGAGA	TGGGAGTCTG	TCTTGCAACA	CCGAGATCGG	GSTGGGCGTC	4800
	AGTGGCTCTT	CATGCTGCTG	CTCTCTGGGA	AAGGCGCTGG	GAACCCCTTG	TGAGACATGC	4860
	CCCTCTGTCA	ATAGCACTGA	ATATTACACC	CTGTGTCCTG	GAGGTGAAGG	CTTCAGACCT	4920
	AAOCCCATCA	CAATCATTTT	AGAAGACATT	GACGAATGCC	AGGAGTTACC	AGGTCTCTGC	4980
70	CAGGTTGGAA	ACTGCAATCA	CACITTTTGG	AGCTTCCAGT	GTGAGTGGCC	ACAAGGCTAC	5040
	TACCTCAGCG	AGGATACCCG	CATCTGTGAG	GATATTGATG	AGTGTTTTGC	ACATCTGCTT	5100
	GTGTGTGGGC	CTGGGACCTG	CTATAACACC	CTGGGAAATT	ACACCTGCAT	TTGCCCACTT	5160
	GAGTACATGC	AGGTCAATGG	AGGCCACAAAC	TGCATGGACA	TGAGAAAAAG	CTTTTGTCTAC	5220
	CGAAGCTATA	ATGGAACCCAC	TTGTGGAAT	GAGTTGCTCT	TCAATGTGAC	AAAAAGGATG	5280
75	TGCTGCTGCA	CATATATGTT	GGGCAAGCT	GGGAACAAAC	CTTGTGAAAC	ATGCCCAACT	5340
	CCAGGAACAG	TGCACTTTAA	AACCATATGT	GGAAATATTC	CTGGATTAC	CTTTGACATT	5400
	CACACAGGAA	AACTGTTTGA	CATTGATGAA	TGTAAGAGAG	TTCCAGGCAT	TTGTGCAAAT	5460
	GGTGTGTGCA	TGCACTTTTC	GGCTGTGAAT	GGCTTACAGG	ATTCAGTTAC		5520
	AATGACCTGC	TGTTGGTTTG	TGAAGATATA	GATGAGTGCA	GCAATGTTGA	TAACTCTCTG	5580
80	CAGCGGAATG	CAGACTGCAT	CAATGATCCT	GGTATGTTACC	GCTGTGAATG	TGCCCGGGT	5640
	TTCAAACTTT	CAOCCAAATG	GGCTGTGTGA	GATGCGAATG	AATGTTTAGA	AATTCCTAAC	5700
	GTTTGCACTC	ATGGCTTGTG	TGTTGATCTG	CAAGGAAGTT	ACCACTGCAT	CTGCCACAT	5760
	GGCTTTAAGG	CTTCTCAGGA	CCAGACCATG	TGCATGGATG	TTGATGAGTG	CGAGCGGCAC	5820
	CCATGTGGAA	ATGGAACCTG	TAAAAACACC	GTGGATCCTT	ATAACTGTCT	GTGCTACCCA	5880
	GGGTTTGAAC	TCACCTATAA	TAAATGATTG	CTGGACATAG	ATGAGTGCAG	TTCTTTTFTT	5940

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70

```

GGTCAGGTGT GCAGAAATGG ACGTTGTTTT AATGAAATTG GTTCTTTCAA GTGCTATGT 6000
AACGAAAGTT ATGAACTTAC CCCAGATGGC AAAAACTGTA TAGACACTAA TGAGTGTGTC 6060
GCCCCCTCCG SCTCTGTGCTC TCCTGGTACC TGTCAGAAAT TGGAGGGATC CTTCAGATGC 6120
ATCTGTGCCC CAGGGTATGA AGTAAAAAGC GAGAACTGCA TTGATATAAA TGAATGTGAT 6180
GAAGATCCCA ACATTGTGCT TTTTGGTTCC TGTACTAATA CTCCAGGGGG CTTCAGTGC 6240
CTCTGCCCCC CTGGCTTTGT ACTATCTGAT AATGGACGGA GATGCTTTGA TACTCGCCAG 6300
AGCTTCTGCT TCACAAATTT TGAATAATGGA AAGTGTCTCT TACCCAAAGC TTTCACACC 6360
ACAAAAGCAA AATGCTGCTG TAGTAAGATG CCAGGAGAGG GCTGGGGGGA CCCCTGTGAG 6420
CTGTGCCCCA AAGACGATGA AGTTGCATTT CAGGATTTGT GTCCATATGG CCATGGAAC 6480
GTCCCTAGTC TTCAATGATC ACGTGAAGAT GTCAATGAGT GTCTTGAGAG CCCAGGCATT 6540
TGTTCAAAAT GTCAATGTAT CAACACCGAC GGATCTTTTC GCTGTGAATG TCCAATGGGC 6600
TACAACTCTG ACTACACTGG AGTACGCTGT GTGGATACTG ATGAGTGTTC AATCGGCAAT 6660
CCGTGTGGAA ATGGTACATG CACCAATGTT ATTGGGAGTT TTGAATGCAA TTGCAATGAA 6720
GGCTTTGAGC CAGGGCCCAT GATGAATTGT GAAGATATCA ACGAATGTGC CCGAACCCTA 6780
CTGCTGTGTG CTTTACGCTG CATGAACACT TTTGGGCTCT ATGAATGCAC GTGCCCGATT 6840
GGCTATGCCC TCAGGGAAGA TCAAAAGATG TGCAAGATC TGGATGAATG TGCTGAAGGG 6900
TTACACGACT TCGAATCTAG GGCATGATG TGTAAAGATC TAATCGGCAC CTTCAGTGC 6960
ATCTGCCCTC CTGGAATGGC CCGAAGGCCG GATGGAGAAG GCTGTGTAGA TGAATAATGA 7020
TTCAGGACCA CAGGAGGAAT CTGTGAAAAT GGACGTTGTG TTAACTTAT TGAAGCTAT 7080
AGATGTGAGT GTAATGAAGG ATTCCAGTCA AGTTCTTCAG GCACTGAATG CCTTGACAA 7140
CGACAGGGTT TCTGCTTTGC AGAGGTACTG CAGACAAATG GTCAATGGC ATCCAGTAGT 7200
CGCAATCTCG TCTGATCTAG AGAATGCTGC TGTGATGGTG GCGAGGCTG GGGCCACCA 7260
TGGAGCTTTT GCCACTTCC TGGAACTGCC CAGTACAAA AGATATGTCC TCATGGCCCA 7320
GGATATACAA TCGTGAAGG AGATATTGAT GAATGTAAAG TAATGCCAAA CCTCTGCAC 7380
AATGGTCACT GCATCAATAC CATGGGCTCA TTCCGATGCT TCTGCAAGT TGGCTACACC 7440
ACGACATCA GTGGAACCTC TTGTATAGAC CTGTGTAATG GCTCCAGTC CCCBAAACCA 7500
TGCACACTCA TCTGCAAGAA CACTGAGGGG AGTTATCAGT GTTCATGTCC GAGGGGGTAT 7560
GTCTGTCAAG AGGATGGAAA GACATGCAAA GACCTTGATG AATGTCAAC AAAGCAGCAT 7620
AATGCCAGT TCCCTGTGTG CACACCCCTG GGGGGGTTTA CTTGTAAATG TCCACCTGCT 7680
TTCACACAGC ATCACACTGC TTGTATCGAC AACACGAAAT GTGGGTCTCA ACCTTTGCTT 7740
TGTGGAGGAA AGGGATCTG TCAAAACACT CCAGGCGATT TCAGCTGTGA ATGCCAAGA 7800
GGGTCTCTCT TGTATGCCAC CCGACTGAAC TGTGAAGATG TTGATGAATG TGATGGGAAC 7860
CACAGGTGCC AACACGGCTG CCAGAACATC CTGGGTGGCT ACAGATGTGG CTGCCCCCAA 7920
GGCTACATCC AGCACAATCA GTGGAATCAG TGTGTGATG AGAATGAATG CTCCAATCCC 7980
AATGCCGTGT GCTCTGCTTC CTGCTACAA ACCTCGGGGA GTTACAAATG CGCCTGCCCC 8040
TGGGGTTCT CTTTGGACCA GTTCTCCAGT GCTTGCCACG ACGTGAATGA GTGCTGTGTC 8100
TCCAGAAACC CTTGCAATTA CGGCTGCTCT AACACGGAGG GGGCTTACCT CTGTGGCTGC 8160
CCCCTGGGT ATTCAGAGT GGGACAAAGC CACTGTGTCT CAGGAATGGG ATTTAAACAG 8220
GGGCAGTACC TGTCACTGGA TACAGAGGTC GATGAGGAAA ATGCTCTGTC CCCAGAAACA 8280
TGCTACAGAG GCAAAATCAA CGGCTATCCT AAGAAAGACA GCAGGCGAAA GAGAAGTATT 8340
CATGAACCTG ATCCCACTGC TGTGAACAG ATCAGCCTAG AGAGTGTGCA CATGGACAGC 8400
CCGTCAACA TGAAGTTCAA CCTCTCCAC CTGCGCTCTA AGGAGCACAT CTTGGAAC 8460
AGGCCCCCA TCCAGCCCTT CAACAACAC ATCGGTTATG TCATCTCTCA AGGGAACGAT 8520
GACAGGTACT TCCGATCCA CCAAGGAAT GGGCTCAGCT ACTTGCAAC GGCCTAAGAAG 8580
AAGCTCATGC CCGCACATA CACACTGGA ATCACTAGCA TCCCTCTCTA CAAGAAGAAG 8640
GAGCTTAAGA AACTGGAAGA GAGCAATGAG GATGACTACC TCTAGGGGA GCTTGGGGAG 8700
GCTCTCAGAA TGAAGCTGCA GATTGAGCTC TATTAACCGT TCACAGACTT GGGCCACAGC 8760
TCAATCTTA GCACAGCCAG TCTGCAGAG CATTTGAAA GTCAAGGACT AATTTTAAAG 8820
AGGAAAATA ATATAACTG TTGTTCTTT CCTCCTGTC TTAGACTTTG AATGTTGACC 8880
CTCACAGGGA GGGATAATTT AGACTCTGCT ATGGCCAAAG ATTGAGCTC AAAGGCAACC 8940
GTGGTACTG TATTTTTAT ATAACCTCAT TTTAAATAT ATTAAAGAA ACCTAAATGT 9000
TCAAGATATC AGCATTTGCG ACTAAATGCA CAAAAATAAT GTGAGCTTTT TTTTTTTTT 9060
CCTGTAGACA GTCTGTAAAC CTTTGGGTAT TTTGCTATAG TTGCTAATTA AAAAAATATA 9120
GATGTTTATT TATTTTAAAT ATGGAGAAAT GAACAACTA TGTAAACAAA 9180
AAGGGAACT CACTTGTTTT TCITTAGATT TATAAATTTG AGCTATTTTT TTTAGAGGTG 9240
CTTTTTAAAT ATCAATAGA TACAAGAGAT GTTTCCTTTG GTTTCTGTC AGTCATCCAG 9300
CIGATACACA CCTGTAGAT TTTAAAGAA GCCACACAGA GCTGAATCG GCACTGTCTA 9360
TCAATAATTT AAAAGACATG AATGTCAATTA GATCCTTTAT AACGTAGATC GAAGCCAAAG 9420
CAGCTCATTT GTGACAACAT TTCATATCAC CAGACACACC AGGCAACAGA AGTTGAAGCA 9480
CAACCACTGT AGCAAAATAC CTTGACTGCT TGTGAGACCA TTAGCATGCG AGGCCAAACC 9540
GTACTGTATT TCCTTCTCAT AACCTCAAGG AACCTATGAT GCTACCACCA ACACTCAT 9600
CTTACCAGAG GTGCGCTGGG TCCTCATGGT ACTGTAGGCA GCTGAAGAAC CGCGTTCCC 9660
TTGAAAGGGA ACACTGACA TTCTGTGGTG TTTCTGCTG TCITTAATAA TGGTGCATT 9720
ATTATGTTCA AGTTATTTCA GGATTGCCAT ATGTGCAAC AAATCATGCA ATGCAGCCAA 9780
GGAAATATAG TTGTGTGTTG TGTTTAAAC CCAITTTTTT TTTAGAAATT TCATTATAC 9840
TGTAGTTATA CACCATATGC CTCATTTTAT CATAGCCTAT TGTGTAGAA AGATGTTTGT 9900
ACAATGAATT GATGTTTAGT TTGCTTAGT CATTTAAAAA GATATTGTAC CAGGATGTGC 9960
TATTAAGAGC ACGTATCCAT TATTTCTCTC AACCCAAAGAA CCTGTTTCTT GGACCAAGTGA 10020
CCAAACCTCA TATGTGAAT GGCCAAAGCA CATGCAGGCT CCGTGTGTTT CCTCTCAAAC 10080
CTGTGCTGAC CAAAGATTAG TAACCAAGTA TACCCAGTAT TTTGAGGTTT TATTTGTTTT 10140
TTAATAACTA AAAAAAACT CAGTCC
  
```

Seq ID NO: 162 Protein sequence
 Protein Accession #: NP_001990.1

75
 80

```

1 11 21 31 41 51
MGRERRLLCLQ LYFLNLGCVV LMAQGTAGQP QPPPKPPRP QPPFQQVRS TAGSEGGFLA 60
PEYREBGA AVSRVRRRQGV DVLGGPNVCG SRFSYCCPG WKTLPGGNQC IVPICRNSG 120
DGFCSRPMNC TCSSQIISST CGSKSIQCS VRMNGGTCA DDHCQCQKBY IGYCQGPVC 180
ENGQNGGRC LAQPCACVYG FTGPQCBEDY RTGPPCTQVM NQWQQQLTG IVCTKTLCA 240
TTGRAMGHPC BMCPAQPPFC RRGFIPNIRT GACQDVDEQ ALPLGICQGN CINTVGSPEC 300
RCPAGHKQSE TTQKCEDIDE CSILFGICET GECSTVGSY FCVCPRGYVT STDGSRCIDQ 360
RTGMCFSLV NGRCAQELFG BMLKMQCCCE PGRCNGIGTI PEACPVRSSE EYRRLCMDGL 420
  
```

5	PMGGIPGSAG	SRPGGTGGNG	FAPSGNGNGY	GPGGTGFIPI	PGNGGFSFV	GGAGVGAGGQ	480
	GPITITGLTIL	NQITIDICKH	AMLCINGRCI	FTVSSYRCEC	NMGYKQDANG	DCIDVDECTS	540
	NPCTNGDCVN	TPGSHYCKCH	AGFQRTPTKQ	ACIDIDBCIQ	NGVLCKNGRC	VNSDGSFQCI	600
	CNAGFELTID	GKNCVDHDEC	ITTNMCLNGM	CINEDGSPKC	ICKPGFVLAP	NGRYCTDVDE	660
	QQTGICMNG	RCINSSGSFR	CDCPGLAVG	MDGRVCVDTH	MRSTCYGSIK	KGVCVRPFPF	720
	AVTKSECCCA	NFDYGPGEPC	QPCPAKNSAE	FHGLCSSGVG	ITVDGRDINE	CALDFDICAN	780
	GICENLRGSY	RCNCNSGYRP	DASGRNCIDI	DECLVNRLLC	DNGLCRNTPG	SYSCTCPFGY	840
	VFTETETICE	DINECESNPC	VNGACRNNLG	SFNCBCSPGS	KLSSTGLICI	DSLKGTCLWN	900
10	IQDSRCEVNI	NGATLKSECC	ATLGAANGSP	CERCELDTAC	FRGLARIKGV	TCEDVNECEV	960
	PFVCPNGRC	VNSKGSFHC	CPEGLTLDGT	GRVCLDIRME	QCYLKWDEDE	CIHPVPKFER	1020
	MDACCCAVGA	AWGTECEBEP	KFGTKEYETL	CPRGAGPANR	GDVLTGRFFY	KDINECKAPP	1080
	GMCTYKCRN	TIGSPKCRN	SQFALDMRER	NCTDIDECRI	SPDLGSGGIC	VNTPGSFECB	1140
	CFEGYESGFM	MMKNQMDIDG	CERNPLLCRG	GTCVNTSSSF	QCDCPLGHEL	SPSREDCVDI	1200
15	NECSLSDNLC	RNGKCVNMIG	TYQCSNPGY	DATPDROGCT	DIDECMIMG	GCDTQCTNSE	1260
	GSYECSCSEG	YALMPDGRSC	ADIDECENNP	DICDGGQCTN	IPGEYRCLCY	DGFMAAMDMD	1320
	TCIDVNECDL	NSNICMFGEC	ENTKGSFICH	QCLGYSVKKG	TTGCTDVDEC	EIGAHNCDMH	1380
	ASCLNIPGSP	KCSGRFVCLG	NGIKCIDLDE	CSNGTBQCSI	NAQCVNTPGS	YRCACSEGFT	1440
	GDGFTCSDDV	ECANENINLCE	NGQCLNVPGA	YRCCEMGFT	PASDSRSQCD	IDECSEFQNIC	1500
20	VSGTCNNLPG	MPHICIDGTY	ELDRTGGNCT	DIDECADPIN	CVNGLCVNTP	GRYECNCPFD	1560
	PQINPTGVGC	VDRVRNGCYL	KPGPRGDSGL	SCNTEIGVGV	SRSSCCCSLG	KAWGNPCZTC	1620
	PFVNSTETTY	LCPGGSGFRP	NPITILIEDI	DEQDELPGLC	QGGNCINTFG	SFQCECPQGY	1680
	YLSEDRTRIC	DIDECFAHPG	VCGPSTCYNT	IGNYTICPP	EYMQVNGGHN	CMDMRKSFY	1740
	REYNATTCEB	ELPFNVTKRM	COCTYNVGA	GNKPCFCEPT	PGTADFTTIC	GNIPGFTFDI	1800
25	HTGKAVDIDE	KCEIPGICAN	GVCINQIGSF	RCECPTGFSY	NDLLLVCEDI	DECSNKNLNC	1860
	QRNADCINSP	GSYRCECAAG	FKLSPNGACV	DRNECLSLPN	VCSEGLCVDL	QGSYQCTCHN	1920
	GFKASQDQTM	CMOVDECERH	PCGNGTCKNT	VGSYNCLCYP	GFEITRNNDI	LDIDECSEFF	1980
	GQVCNMGRCF	NEIGSRFVCL	NBYELTPDG	KNCIDTNECV	ALPGSCSPGT	CQMLEGSEFC	2040
	ICPPGYEVKS	ENCIDINECD	EDFNICLFSS	CINTPGGQPC	LCPGPFVLSO	NGRRCFDTRQ	2100
30	SFCPTNFENG	KCSVPKAFNT	TKAKCCCSKM	POEGWGDPCF	LCPKDEVAFF	QDLCPYGHGT	2160
	VPSLHDTRED	VNECLEBPGI	CSNGQCINTD	GSFRCECPMG	YNLDYTGVR	VDTDECSIGN	2220
	PCGNGTCTNV	IGSFECNCNE	GFEPGEMMNC	EDINECAQNP	LLCALRCMNT	FGSYECTCPI	2280
	GYALREDQRM	CKDIDECABG	LEDCESRGMM	CKNLIGTFMC	ICPPGMARRP	DGSGCVDENE	2340
	CRTPGICEN	GRCVNIIGSY	RCEBNEGFQS	SSSGTECLDN	RQGLCPAEVL	QTICQMASSS	2400
35	RNLVTKSECC	CDGGRGNHGQ	CELCLPLGTA	QYKKICPHGP	GYTIDGRDID	ECKVMPNLCT	2460
	NGQCINTMGS	FRCECKVGYT	TDLSGTSCID	LDCECSQSPK	CNYICKNTEG	SYQCSCPRGY	2520
	VLQEDGKTCK	DLDEGQTKQH	NQQLFCVNTL	GGFTCKCPFG	FTQBTHTACID	NNECGSQPLL	2580
	CGGKIGCQMT	PGFSFSCQJR	GFSLDATGLN	CEVDDECDGN	HRCQHGQONI	LGGRYCGCPQ	2640
	GYIQHYQWNG	CVDNEBENBP	NACGSASCVN	TLGSYKCAFP	SGFSFDQFSS	ACHDVNECSS	2700
40	SKNPNYNGCS	NTEGGYLCQG	PPGYRVGQG	HCVSGMGFNK	GOYLSLDETV	DEENALSPEA	2760
	CYECKINGYP	KKDSRQKRBI	HEPDPTAVEQ	ISLESVDMDS	PVMKFLNLSH	LGSKEHILEL	2820
	RPAIQPLNMH	IRYVISQGNQ	DSVFRTHQRN	GLSYLHTAKK	KLMPGTYYLS	ITSIPLYKKK	2880
	ELKKLEESME	DDYLLGLLGE	ALRMRLQIQ	Y			
45	Seq ID NO: 163 DNA sequence						
	Nucleic Acid Accession #: NM_013372.1						
	Coding sequence: 63..617						
50	1	11	21	31	41	51	
	GGGGCCGCAC	TCAGCGCCAC	GCGTCGAAAG	CGCAGGCCCC	GAGGACCCGC	CGCACTGACA	60
	GTATGAGCGG	CACAGCCTAC	ACGGTGGGAG	CCCTGCTTCT	CCTCTGGGGG	ACCGTGTCTG	120
	CGGTGTCTGA	AGGGAAGAAAG	AAAGGGTCCC	AAGGTGCCAT	CCCCCGCCCA	GACAAAGGCC	180
	AGCACAATGA	CTCAGAGCAG	ACTCAGTCCG	CCCAGCAGCC	TGGCTCCAGG	AACCGGGGGC	240
55	GGGGCCAGAG	CCGCGGCACT	GCCATGCCCG	GGGAGGAGGT	GCTGGAGTGC	AGCCAAAGAG	300
	CCCTGCATGT	GACGGAGCGC	AAATACCTGA	AGCGAGACTG	GTGCAAAACC	CAGCCGCTTA	360
	AGCAGACCAT	CCAAGAGGAA	GGCTGCAACA	GTGCGACCAT	CATCAACCGC	TTCTGTATAG	420
	GCCAGTGCAA	CTCTTTCTAC	ATCCCCAGGC	ACATCCGGAA	GGAGGAAGGT	TCCTTTCAGT	480
	CTCTCTCTCT	CTGCAAGCCC	AAGAAATTC	CTACCATGAT	GGTCACTCTC	AATCTGCCCTG	540
60	AACATACAGC	ACCTACCAAG	AAGAAGAGAG	TCACACGTGT	GAAGCAGTGT	CGTTGCATAT	600
	CCATCGATT	GGATTAAGCC	AAATCCAGGT	GCACCCAGCA	TGTCTTAGGA	ATGCAGCCCC	660
	AGGAAGTCC	AGACCTAAAA	CAACCGATT	CTTACTTGGC	TTAAACCTAG	AGGCCAGAG	720
	AACCCCCAGC	TGCTTCTTGG	CAGGAGCCTG	CTTGTGCGTA	GTTCGTGTGC	ATGAGTGTGG	780
	ATGGGTGCTT	GTGGGTGTTT	TTAGACACCA	GAGAAACAC	AGTCTCTGCT	AGAGAGCACT	840
65	CCCTATTGTT	TAAACATATC	TGCTTFAATG	GGGATGTACC	AGAAACCCAC	CTCACCCCGG	900
	CTCACATCTA	AAGGGGCGGG	GCGGTGGTCT	GGTCTGACT	TTGTGTTTT	GTGCCCTCCT	960
	GGGGACCAGA	ATCTCCTTTC	GGATGTAAATG	TTCTATGAGG	AGGCTCTCTCT	GAGGGCAGGA	1020
	GACCTGTTTT	AGTGTCTGCAT	TCGACATGGA	AAAGTCTCTT	TAACTGTGTC	TTGCATCTCT	1080
	CTTCTCTCT	CCTCTCTACA	ATCCATCTCT	TTCTTAAGTGT	ATAGTGACTA	TGTGAGTCTA	1140
70	ATCTCTGTT	TGCCAAGGTT	CCTAAATTAA	TTCACTTAAC	CATGATGCAA	ATGTTTTCAT	1200
	TTTTTGAAG	ACCCCTCAGA	CTCTGGGAGA	GGCTGGGTGT	GGCAAGGACA	AGCAGGATAG	1260
	TGGAGTGAAG	AAGGGAGGGT	GGAGGGTGGG	GCCAAATCAG	GTCCAGCAAA	AGTCAGTAGG	1320
	GACATTGCGA	AGCTTGTAA	GGCCAAATACC	AGAACACAGG	CTGATGCTTC	TGAGAAAGTC	1380
	TTTTCTTAGT	ATTTAAACAGA	ACCCAAAGTGA	ACAGAGGAGA	AATGAGATTG	CCAGAAAGTG	1440
75	ATTAACTTGT	AGCGTGTGAA	TCTGCTCAAA	CCTAACACCA	AACAGAAAC	ATAAATACGT	1500
	ACCACTCTTA	TGTTGGGACC	CNAGCAAGTT	AGCTAAACCA	AACCAACTCC	TCTGCTTGT	1560
	CCCTCAGGTG	GAAAAGAGAG	GTAGTTTAGA	ACTCTCTGCA	TAGGGGTGGG	AATTAATCAA	1620
	AAACCKCAGA	GGCTGAAATT	CCTAATACCT	TTCTTTATC	GTGGTTATAG	TCAGCTCTCT	1680
	TCCATTCAC	TATTTCCCAT	AATGCTCTCG	AGAGCCACTA	ACTTGATTGA	TAAAGATCCT	1740
80	GCCTCTGCTG	AGTGTACATG	ACAGTAAGTC	TAAAGATGAG	AGAGTGTAGG	GACTACTCTG	1800
	TTTTAGCAAG	ARATATTCTG	GGGGTCTTTT	TGTTTAACT	ATTGTGAGGA	GATTGGGCTA	1860
	RAGAGAAAG	GACGAGAGTA	AGGAAATAAA	GGGATTTGCC	TCTGGCTAGA	GAGTAAGTAA	1920
	GGTGTAAATA	CCTGGTAGAA	ATGTAAAGGA	TATGAACCTC	CTTTCTTTAT	GTGCTCACTG	1980
	AGGATCTGAG	GGGACCTGT	TAGGAGAGCA	TAGCATCATG	ATGTATTAGC	TGTTCACTCTG	2040
	CTACTGGTGG	GATGACATA	ACTATTGTAA	CTATTGAGTA	TTTACTGGTA	GGCACTGTCC	2100

TCTGATTAAA CTTGGCCTAC TGGCAATGGC TACTTAGGAT TGATCTAAGG GCCAAAGTGC 2160
 AGGGTGGGTG AACITTTATTG TACTTTGGAT TGGTTAACC TGTITTTCTT AAGCCTGAGG 2220
 TTTTATATAC AAACCTCCCTG AATFACCTTT TTGCCTTGTA TCTTCTCAGC CTCCTAGCCA 2280
 AGTCCATAGT AATATGGAAA ACAACACTG CAGACTTGAG ATTCAAGTGC CGATCAAGGC 2340
 TCTGGCATTG AGAGAACCCCT TGCAACTCGA GAAGCTGTGT TTATTTCTGT TTGTITTTGA 2400
 TCCAGTGTCT TCCCATCTAA CAACTAAACA GCAGCCATTT CAAGGCGGGA GATATTTTAA 2460
 ACACCCAAAA TGTGGGTCTT GATTTTCAAA CTFTTAAACT CACTACTGAT GATTCTCAGC 2520
 CTAGGCGAAT TTGTCCAAAC ACATAGTGTG TGTGTTTGT ATACACTGTA TGACCCCAAC 2580
 CCAATCTTTT GTATTGTCCA CATTCTCCAA CAATAAAGCA CAGAGTGGAT TTAATTAAAG 2640
 ACACAAATGC TAAGGCAGAA TTTTGAGGGT GGGAGAGAAG AAAAGGGAAA GAAGCTGAAA 2700
 ATGTAAAACC ACACGAGGGA GGAAAATGA CATTCAAGAC CAGCAAAACAC TGAATTTCTC 2760
 TGTGTGTTTT AACTCTGCCA CAGAATGCA ATTTGTTTAA TGGAGATGAC TTAAGTTGGC 2820
 AGCAGTAATC TTCTTTTAGG AGCTGTATCC ACAGTCTTGC ACATAAGTGC AGATTTTGGC 2880
 CAAGTAAAGA GAATTTCCCT AACACTAACT TCACITGGAT AATCAGCAGC GTAACTACCC 2940
 TAAAGCATA TCACTAGCCA AAGAGGGAAA TATCTGTCT TCTACTGTG CCTATATTAA 3000
 GACTAGTACA AATGTGTGT GTCTTCCAAC TTTCATTGAA AATGCCATAT CTATACCAAT 3060
 TTTTATTGCA GTCACTGATG ATGTAATGAT ATATTTTTC ATTATATAG TAGAATATT 3120
 TTATGGCAAG ATATTTGTGG TCTTGATCAT ACCTATTAAA ATAATGCCAA ACACCAATA 3180
 TGAATTTTAT GTATCTTGG TGTGTCTGG CATTAAAAGA AAAAACAACA CATCTGGAA 3240
 GTCTGTAAGT TGTTTTGTG TACTGTAGGT CTTCAAAGTT AAGAGTGTAA GTGAAAAATC 3300
 TGGAGGAGAG GATAATTTCC ACTGTGTGGA ATGTGAATAG TTAATGAAA AGTTATGGTT 3360
 ATTTAATGTA ATTATTACTT CAAATCCITT GGTCACTGTG ATTTCAAGCA TGTTTTCTTT 3420
 TTCCTCTTTA TATGACTTTC TCTGAGTTGG GCAAGAAGA AGCTGACACA CGTATGTTG 3480
 TTAGAGTCTT TATCTGTCT AACCTAACTA AAATCTTGAC CCAGCTGAAC ATGTCTTCTT 3540
 GAGTCAAGTC CTGAATCTTT ATTTTFAAA TTGAATGTTT CTTAAGGTTT AACATTTCTA 3600
 AAGCAATATT AAGAAGACT TTAAATGTTA TTTTGAAGA CTTACGATGC ATGTATACAA 3660
 ACGAATAGCA GATAATGATG ACTAGTTCAC ACATAAAGTC CTTTAAAGGA GAAAATCTAA 3720
 AATGAAAAGT GGAATAACAG AACATTITATA AGTGATCAGT TAATGCCATA GAGTGAAAGT 3780
 AGTTCTATTG ACATTCTCTA AGATATTTAA TATCAACTGC ATTATGTATT ATGTCTGCTT 3840
 AAATCATTTA AAACCGCAA AGAATTATAT AGACTATGAG GTACCTTGCT GTGTAGGAGG 3900
 ATGAAGGGG AGTTGATAGT CTCATAAACC TAATTTGGCT TCAAGTTTCA TGAATCTGTA 3960
 ACTAGAATTT AATTTTCAAC CCAATATATG TCTATATAGC CTTTGCTAAA GAGCAACTAA 4020
 TAAATTAAAC CTATCTTTC AAAAAAAA

Seq ID NO: 164 Protein sequence
 Protein Accession #: NP_037504.1

1 11 21 31 41 51
 MSRTAYTVGA LLLLLGTLFP AAEKKKKGSQ GAIPFPDKAQ HNDSEQTQSP QQFGSENRGR 60
 GQGRGTAMPG EEVLSSQSEA LHWTERKYLK RDWCKTQPLK QTTHBECNRS RTIINRFQYB 120
 QQNSPYIPRH IRKEBGSFQS CSFCKPKKFT TMMVTLNCPK LQPPTKIKRV TRVKQCRCLB 180
 IDLD

Seq ID NO: 165 DNA sequence
 Nucleic Acid Accession #: CAT cluster

1 11 21 31 41 51
 GAATTGCATC GGACAGAGCT TCGCCATGGC CBAAGTTGCGC GTGGTTGTGG CACTAACACT 60
 GCTACGTTTC CGCTGAGCG TGGACCGAAC CGCCAGGTG CCGCCGAGC CGGAGTCTAT 120
 ACTGCGCACG GAGAACGGGC TCTGGCTCAA G

Seq ID NO: 166 DNA sequence
 Nucleic Acid Accession #: FGENESH predicted
 Coding sequence: 1..1650

1 11 21 31 41 51
 ATGCCACCTC TGCCATCCAG AATGTCTCGG GGGCCCTCCT TGGTGACAGG CAGGATGCTG 60
 CCCATCAGAG ACCGCTGTCT GCACCTCTCG GGGCTGGAGA AGACGGCGTT CGCATATAC 120
 GGGGTGTCCA CCGTCTCTCT CTCTCTCTCT TCTTCTCTGT TCCGCTCTGT GCTGCGTTC 180
 CTGAGGCTCT GCAGGAGCTT CTACATCACT TGCCTGCGGC TCGCTGCTCT CCGCCAGCCT 240
 CCGCGGCGCA ACTGGCTGCT GGGCCACCTG GGCATGTACC TTCCAAATGA GCGGGCCCTT 300
 CAAGATGAGA AGAAGGTACT GGACAACATG CACCATGTAC TCTTGGTATG GATGGGACCT 360
 GTCTGCGGCG TGTGTGTTCT GGTGACCTCT GATTACATCA AACCCCTTTT GGGAGCTTCA 420
 GCTGCGATCG CCGCCAAAGG TGACCTCTTC TATGGCTTCC TAAACCTTGT GCTAGGGGAT 480
 GGGCTGCTGC TCAGCAAGGG TGACAAGTGG AGCCGSCAOC GTGCTCTGCT GACACCGGCC 540
 TTCCACTTTG ACATCTGAA GCGTTACATG AAGATCTTCA ACCAGAGCGC TGACATTATG 600
 CATGCTAAT GGCGGCTCT GGCAGAGGGC TCAGCGTCT CCGTTGATAT GTTTGAGCAT 660
 ATCAGCCTCA TGACCTGGA CAGTCTTCAG AAATGTGTCT TCGCTACAA CAGCAACTGC 720
 CAAGAGAGA TGAATGATTA TATCTCGGCT ATCATTGAAC TGAGCGCTCT GTCTGTCTCG 780
 CGCCAGTATC GCTTGCACCA CTACCTCGAC TTCTTTACT ACCGCTCGCT GGATGGGCGG 840
 AGGTTCCGGC AGGCTGTGTA CATGGTGCAC CACTTCACCA CTGAAGTCAT CCAGGAACGG 900
 CGCGGGCAC TGCCTCAGCA GGGGGCCGAG GCTTGGCTTA AGGCCAAGCA GGGGAAGACC 960
 TTGGACTTTA TTGATGTGCT GCTCTCTGCC AGGGATGAAG ATGGAAGGA ACTGTCAGAC 1020
 GAGGATATCC GAGCCGAAGC AGACACCTTC ATGATTGAGG GTCAACGAC AACATCCAGT 1080
 GGGATCTCTT GATGTCTGTT CAATTGGCA AAGTATCCGG AATACCAGGA GAATGCGGGA 1140
 GAGAGATTTC AGGAAGTCAT GAAAGGCGCG GAGCTGGAGG AGCTGGAGTG GGAAGATCTG 1200
 ACTCAGCTGC CTTTACAAC TATGTGCATT AAGGAGAGCC TGCGCGATA CCACTCTGTC 1260
 ACTCTGTCT CTCGCCAATG CAGGAGGAC ATCAAGCTCC CAGATGGGCG CATCATCCCC 1320
 AARGGAATCA TCTGCTTGGT CAGCATCTAT GGAACCCAC ACACCCACAG AGTGTGGCCT 1380
 GACTCCAAGG TGTACAACCC CTACCGCTTT GACCGGACA ACCCAACAGA GCGCTCTCCA 1440

CTGGCCTATG TGCCCTCTC TGCAGGACCC AGGAATTGCA TCGGACAGAG CTTCGCCATG 1500
 GCCGAGTTGC GCGTGGTTGT GGCCTAACA CTGCTACGTT TCCGCTGAG CGTGGACCGA 1560
 ACGCGCAAGG TCGCGCGGAA GCGCGAGCTC ATACTGCGCA CGGAGAACGG GCTCTGGCTC 1620
 AAGGTGGAGC CGCTGCCTCC GCGGCGCTGA

Seq ID NO: 167 Protein sequence
 Protein Accession #: FGENESH predicted

1 11 21 31 41 51
 | | | | |
 MFPLPSRMSR GPSLVTGRML PITDRLLHL GLEKTAFRY AVSTILLFL FFLFRILLRF 60
 LRLCRSFYIT CRRLCRFPQ PRNRWLLGHL GMYLPNEAGL QDEKKVLDNM HRVLLVMMGP 120
 VLPFLVLVHP DYIKPLLGAS AAIAPKDDLF YGFLKPHLED GLLLSRGDKW SRHRLLTPA 180
 FHFIDILKPYM KLFNQSDIM HAKWRHLAEG SAVSLDMFEH ISLMTLDSLQ KCVFSYNSNC 240
 QEKMSDYISA IIELSALSVR RQYRLHHYLD FYVRSADGR RFRQACDMVE HFTTEVIQER 300
 RRALRQGGAE AMLKAKQGKT LDFIDVLLA RDEDGKELSD EDIRAEADTF MFEGHDTTSS 360
 GISWMLFNLA KYREYQSKCR EEIQEVMKOR ELELEWDDL TQLPFTTMC I KESLRQYPPV 420
 TLVSRQCTED IKLPDGRITP KGIICLVSIY GTHNPVTWP DSKVYNPYRF DPDNPQQRSP 480
 LAXVFPFAGP RNCIGQSFM ABLRVVVALT LLRFRLSVDR TRKVRKRPKL ILRTENGLWL 540
 KVEFLPFRA

Seq ID NO: 168 DNA sequence
 Nucleic Acid Accession #: AK058088.1
 Coding sequence: 252..1772

1 11 21 31 41 51
 | | | | |
 AGGAAACCAA GCGAAGCTCC CCTGTCAAA GCACCTTGGC CCATAAGAAG AAAAGGGGGA 60
 GCCCCAGATG TGATGAGCGC TTCCAGGCTT CAGGCTCAGA AGGCGCCGCC AGCTCTCTCT 120
 TAACCTCAGAG GCCAGTGTGA TGGGAGTTCC TCCACTCAGC AACTTCCGCC TGTAAACACG 180
 CCGTGGGTGG GCAAAAGGCG TTGGGAACGG TTGCTGTCT TTTCTCTCCT GCGTAATTTT 240
 CACTTTCATT CATGATAATG TCGAACACGC ACAAAGCTCG GCTGGAACGC CGGCTCACTG 300
 GCTCAACCAA CCGTGTGGGT TTGCCCAAAC AGCTTTTCTC TGGGACCTCG CTCTCACTTT 360
 CCCAGATGTG CAGGCTCTG AGCATAGACT TTGAGGAAGC TTGAGGAAC CCAGACAGGT 420
 TATGCATTTT ACATAATCCAG AAGTTTTTCT TTGAGAAATT CAAGAACACG GACATCCAAA 480
 GTGGGGAAGC AGATGTGATT CTGAGTGCC TGGGCTTCAA ATGGGAGCTC CATCAGCCCC 540
 GGCTTTTTC A GTCGTAGACC TTGGCCAAGC TCTACCTGAA AGCCCTGGCG CAGGGCACCA 600
 CACACCCCCC GAGGGAGCTG GAGGAGCTTC TCGAGCTCA ATCAGCTAAG AAGACCAAG 660
 AAAAATCCCC TGCAAGAGCG ATCATCATTT CCTGAAGAT CAATGACCCA CTGGTCACTA 720
 AAGTCGCCCT CCGCAAGGCG CTGAGAACCT TCTACATGAG TGAGGTGGAG ATTAACCTTG 780
 AAGACCTACT GGGAGTGGTG GCTTCCGCCC ACATCCTCCA GTTCAGTGGC CTGTTTCAA 840
 GGTGGGTGGA TGTGATGATA GCCAGACTCA AGCCAGCAC CATCAAGAAA TTCTACGAGG 900
 CCGGCTGCAA GTACAGGAA GAGCAGCTCA CACCCGCTG CGAGAGTGG CTGGAATGA 960
 ACTTGGITTC TCTAGCGGCG ACACAGATCC ACCTCCACAA AATCCACAG GACCTGCTCC 1020
 ACAAAAGTCT GAAGTCCCCC AGGTATTAT CTTTAGTGA ATTCCATCT CTGAAAACAA 1080
 TGTCTTTGTG GGTCTTCTG CAACGAACT ACAAGATTCA GGCAATTCG ACTTATGAAA 1140
 CGGTGATGAC ATTTTTAAAG AGCTTTCTGT AGAAGCTGTG CTTTCTGGAC CGGACATAG 1200
 SAOGGAGCTT GAGGCGCTC TTCTCTGTCT TCGTCTGCA CGGCATCACC AAGGCAAGG 1260
 ATCTGGAGGT GCTGCGGCAC CTTAACTTCT TCCAGAGTCC ATGGCTCGAC CAGGTTACAG 1320
 TCAACCATTA CCAAGCAGTG GAGAAATGGG GGCACATGAT CCACCTGAAA GATCTTAACA 1380
 CCGAGCTGT GAGATTGGG CTGCTCTTAA ACCAGAGAAA TACAACCTAT TCGAAAACGA 1440
 TTGCTCTATA TGGATTCTTC TTTAAGATAA AGGGACTCAA ACATGATACT ACCTCTATA 1500
 GTTTTTACAT CGAGAGAATA AAGCACACAG ACCTGAATC TCCCTCTGCG GTCTACGAGC 1560
 ACACCAAGCT CAGCTTGGCA GCGCAAGCC TGGTGAAGTA TGAGATCAGA GCAGAGGCC 1620
 TGGTGAAGCG CAGAGTGGCG GAGTTCAGGA CAACACAGAT CAAGCAGAG TTTGGGTTGA 1680
 CCACTCATC CTGCAAAAGC CATACCTTGA AATCCAAAC TGTGGGCATC CCAATCTATG 1740
 TAAATTTTGC ATTCACTTTC CCAGCATCTT GACAGTTTCC AGAAGAATCT ATGGGATTTT 1800
 CCCCCACTG GTCTGCATAA AAGAAAATAA AATGACATAA AAGGGAGC

Seq ID NO: 169 Protein sequence
 Protein Accession #: BAB71658.1

1 11 21 31 41 51
 | | | | |
 MIMSNTHKAR LERRVTGSTN RWRLPKQPPS GDLISLSQMC KALSIDFREA LRNFDRLCIS 60
 QIQKFFPENF KKKDIQSGEA DVILECLGFK WELHQPRLFQ SETLAKLYLK ALAQGTTEPL 120
 RELEELLRAQ SPKKTKEKSP AKRIIISLKI MDPLVTKVAF ATALKNLYMS EVEINLEDLL 180
 GVLASAHILQ FSGLFQRCVD VMIAHLKPSI IKKPYBAGCK YKESQLTTGC EKWLENNLVP 240
 LGGTQIHLHX IPQDILHKL KSPRLFTFSE PHLLKIMLLW VFLQLNXYKIQ AIPTYETVMT 300
 FFKSFPENCC FLDRDIGRSL RFLFLCLRHL GITKKGDLV LRHLNFFPES WLDQVTVMHY 360
 HALENGGDMV HLEDLNTQAV RFGLLENQEN TTYSKFLALY GFFPKIKGLK HDITSYSFYM 420
 QRIKHFDLES PSAVYEHMHV SLRAARLVKY BIRAEALVDG KWQEFRTNQI RQKPLGTTSS 480
 CKSHTLKIQT VGPIYVSFA PIPPAS

Seq ID NO: 170 DNA sequence
 Nucleic Acid Accession #: NM_007000.1
 Coding sequence: 1...777

1 11 21 31 41 51
 | | | | |
 ATGCGGTCTG CGGACGACGC GGAGGCCGAG AAGGGATCTC CAGTTGTGGT GGGCCTGCTA 60
 GTTGTGGGCA ATATCATATAT TCTGCTGTCA GGCTGTGCCC TGTGTGCTGA GACCATATGG 120
 GTGACAGCG ACCAGTACCG TGTATACCCA CTGATGGGAG TCTCAGGCAA GGATGACGTC 180
 TTGCTGGTG CCTGGATTGC CATCTTCTGC GGCTTCTCTT TCTTCAATGGT AGCCAGTTTT 240

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

```

GGTGTGGGTG CCGCACCTCG CGGCGCGCGG TCCATGGTCC TCACGTACCT GGTGCTCATG 300
CTCATCGTCT ACATCTTCTGA GTGCGCCTCC TGCATCACGT CCTACACCCA CCGTGACTAC 360
ATGGTGTCCA ACCCATCCCT GATCACCAAG CAGATGCTGA CCTTCTACAG CCGGACACCC 420
GACCAGGGCC AGGAGCTGAC CGGCTCTGG GACCGCTCA TGATTGAGCA AGAATGCTGT 480
GGCACATCTG GTCCCATGGA CTGGGTGAAC TTCAGCTCAG CCTCCCGGC GGCACTCCG 540
GAGGTGGTGT TCCCTCTGCC CCACTGTGC TGTGCGCGGA CGGGAACAT CATCCCCCTC 600
AACGAGGAGG GCTGCGCCTT GGGGCACATG GACTACCTGT TCACCAAGGG CTGCTTCGAA 660
CACATCGGCC ACGCCATCGA CAGCTACACG TGGGGTATCT CGTGGTTTGG GTTTCCTATC 720
CTGATGTGGA CGCTCCCGGT CATGCTGATA GCCATGTATT TCTACACCAT GCTCTGA

```

Seq ID NO: 171 Protein sequence
Protein Accession #: NP_008931.1

1 11 21 31 41 51
MASAAAAEAE KSSPVVVGLE VVGNIIILS GLSLFAETIN VTADQYRVYP LMGVSGKDDV 60
PAGAWIAIFC GFSEFFMVASP GVGAALCRRR SMVLTYLVLN LIVYIFECAS CITSYTHRDY 120
MVSNPFLITK QMLTFYSADT DQGQELTRIN DRVMIEQECG GTGGPMDWVN FTSAFRAATP 180
EVVFPWPELC CRRTGNFIPL NEBGCRLGEM DYLFTRKCFE HIGHALDSYT WGISWFGFAI 240
LMWTLFVMLI AMYFYTM

Seq ID NO: 172 DNA sequence
Nucleic Acid Accession #: NM_006760.1
Coding sequence: 39..593

1 11 21 31 41 51
GAAAGCCCTG CAGCACCTAT TCCACCTCCC AGCCAGCAT GGCACCCCTG CTGCCCATCC 60
GGACCTTGCC CTGTATCCTG ATTCTGCTGG CTCTGCTGTC CCCAGGGGCT GCAGACTTCA 120
ACATCTCAAG CCTCTCTGCT CTGCTGTCCC CGGCGCTAAC GGAGAGCCTG CTGGTGTGCT 180
TGCCCCCCTG TCACCTCACA GGAGGCAATG CCACACTGAT GGTCCCGAGA GCCAATGACA 240
GCAAGTGTGT GACGTCCAGC TTTGTGGTGC CTCOGTGGCG TGGGCGCAGG GAACCTGGTA 300
GTGTGGTGGG CAGTGTGTCT GGCTTCACAG TCACTCGGCT CAGTGCATAC CAGGTGACAA 360
ACCTCGTGCC AGGAACCAAA TTCTACATT CTACCTAGT GAAGAAGGGG ACAGCCACTG 420
AGTCCAGCAG AGAGATCCCA ATGTCCACAC TCCCTCGAAG GAACATGGAA TCCATTGGGC 480
TGGGTATGGC CGCACAGGG GGCATGGTGG TCATCACGGT GCTGCTCTCT GTGCCCATGT 540
TCTGTCTGGT GCTGGGCTTC ATCATTGCCC TGGCACTGGG CTCCCGCAAG TAAGGAGGTC 600
TGCCCCGAGC AGCAGCTTCT CCAGGAAGCC CAGGGCACCA TCCAGCTCCC CAGCCACCT 660
GCTCCAGGC CCCAGGCTG TGGCTCCCTT GGTGCGCTCG CCTCTCTCTC CTGCGCTCTC 720
CTCCCCAGA GCCCTCTCT CCTCTGTCC CTCTCTCTGC CCCCAGTGCC TCACTCTCCA 780
ACACTCCATT ATCTCTCTCA CCCCCTCTCT GTGAGATTG ACTTTCTCTC CATTTTACCA 840
CTTTAAACAC CCCCATAACA ATTCGCCCAT CTTTCTAGTA ACTAAGTCCC TATAATAAAG 900
GCTGAGGCTG CATCTGCCAA AAAAAAAAAA AA

Seq ID NO: 173 Protein sequence
Protein Accession #: NP_006751.1

1 11 21 31 41 51
MAPLLPRTL PLILLALL SPGAADFNIS SLGILLSPAL TESLLVALPF CHLTGGNATL 60
MVRANDSKV VTSSFFVPPC RGRRLVSVV DSGAGFTVTR LSAYQVINLV PGTKFYISYL 120
VKKGTATESS REIPMSTLPR RMBSIGLGM ARTGGMVVIT VLLSVAMPLL VLGFIIALAL 180
GSRK

Seq ID NO: 174 DNA sequence
Nucleic Acid Accession #: Bos sequence
Coding sequence: 1..2733

1 11 21 31 41 51
ATGAAAGTTG GAGTGTCTG GCTCATTTCT TTCTCACCT TCACTGACGG CCACGGTGGC 60
TTCTGGGGGA AAAATGATGG CATCAAAACA AAAAAAGAAC TCATTGTGAA TAAGAAAAAA 120
CATCTAGGCC CAGTCCGAAG ATATCAGCTG CTGCTTCAGG TGACCTATAG AGATTCCAG 180
GAGAAAAGAG ATTTGAGAAA TTTTCTGAAG CTCTTGAAGC CTCCATTATT ATGGTCCAT 240
GGGCTAATTA GAATTATCAG AGCAAAAGGCT ACCACAGACT GCAACAGCCT GAATGGAGTC 300
CTGCAGTGA CCTGTGAAGA CAGCTACACC TGGTTTCTCT CCTCATGCT TGAATCCCGC 360
AACTGCTACC TTCACACGGC TGGAGCACTC CCAAGCTGTG AATGTCACT CAACAACCTC 420
AGCCAGAGTG TCAATTTCTG TGAGAGAAACA AAGATTGGGG GCACTTTCAA AATTAAATGA 480
AGGTTTACAA ATGACCTTTT GAATTCATCT TCTGCTATAT ACTCCAAATA TGCAAATGGA 540
ATTGAAATTC AACTTAAAAA AGCATATGAA AGAATTCAAG GTTTTGAGTC GGTTCAAGTC 600
ACCCAATTTC GAAATGGAAG CATGTTGCT GGGTATGAAG TTGTTGGCTC CAGCAGTGCA 660
TCTGAACTGC TGTACGCCAT TGAACATGTT GCGAGAGAGG CTAAGACAGC CCTTCACAG 720
CTGTTTCCAT TAGAAGACGG CTCTTTGAGA GTGTTGCGAA AAGCCAGTGT TAATGACATT 780
GTCTTTGGAT TTGGGTCCAA GGATGATGAA TATACCTGCG CTTGACAGAG TGGCTACAG 840
GGAAACATCA CAGCCAAGTG TGAGTCCCTT GGGTGGCAGG TCATCAGGGA GACTTGTGTG 900
CTCTCTCTGC TTGAAGRACT GAACAAGAAAT TTCAGTATGA TTGTAGGCAA TGCCACTGAG 960
GCAGCTGTGT CATCTCTGCT GCAAAATCTT TCTGTCACTA TTGGGCAAAA OCCATCAACC 1020
ACAGTGGGGA ATCTGGCTTC GGTGGTGTGC ATTCTGAGCA ATATTTCATC TCTGTCACTG 1080
GCCAGCCATT TCAGGGTCTC CAATTCAACA ATGGAGGATG TCATCAGTAT AGCTGACAAT 1140
ATCCTTAAT CAGGCTCAGT AACCAACTGG ACAGTCTTAC TGGGGGAAGA AAGTATGCC 1200
AGCTCAGGT TACTAGAGAC ATTAGAAAAC ATCAGCACTC TGGTGCCTCC GACAGCTCTT 1260
CCTCTGAAT TTTCTGGGAA ATTCATTGAC TGGAAAAGGA TTCCAGTGAA CAAAAGCCAA 1320
CTCAAAGGG GTTACAGCTA TCAGATTAAA ATGTGTCCCG AAAATACATC TATTCCCATC 1380
AGAGGCCGTG TGTTAATTGG GTCAGACCAA TTCCAGAGAT CCTTCCGAA AACTATTATC 1440

5	AGCATGGCCT CGTGTACTCT GGGGAACATT CTACCCGTTT CCAAAATGG AAATGCTCAG 1500
	GTCAATGGAC CTGTGATATC CACGGTTATT CAAAACTATT CCATAAATGA AGTTTTCCTA 1560
	TTTTTTTCCA AGATAGAGTC AAACCTGAGC CAGCCTCATT GTGTGTTTGG GGATTTCAGT 1620
	CAITTTGCACT GGAAACBAGC AGGCTGCCAC CTAGTGAATG AAACCTCAGA CATCGTGACG 1680
	TGCCAATGTA CTCACCTGAC CTCCTTCTCC ATATTGATGT CACCTTTTGT CCCCTCTACA 1740
	ATCTTCCCGG TTGTAAATG GATCACCTAT GTGGGACTGG GTATCTCCAT TGGAACTCTC 1800
	ATTTTATGCC TGATCATCGA GGCTTTGTIT TGGAAACAGA TTAATAAAG CCAAACTCT 1860
	CACACAGTCC GTATTTGCACT GGTGAACATA GCCCTGTCCC TCTTGATTGC TGATGCTGG 1920
10	TTTATTGTTG GTGCCACAGT GGACAACACG GTGAACCCCT CTGGAGTCTG CACAGCTGCT 1980
	GTGTTCTTTA CACACTTCTT CTACCTCTCT TTGTTCTTCT GGATGCTCAT GCTTGGCATC 2040
	CTGCTGGCTT ACCGGATCAT CCTCGTGTC CATCACATGG CCCAGCATTT GATGATGCT 2100
	GTGGGATTTT GCCTGGGTTA TGGGTGCCCT CTCATTATAT CTGTCAATAC CATTGCTGTC 2160
15	ACGCAACCTA GCAATACCTA CAAAAGGAAA GATGTGTGTT GGCTTAACIG GTCCAATGGA 2220
	AGCAAAACCC TCTTGGCTTT TGTTGTCCCT GCACCTGGCTA TTGCGCTGT GAACCTCGTT 2280
	GTGGTCTGCG TACTTCTCAC AAAGCTCTGG AGGCGGACTG TTGGGGAAAG ACTGAGTCGG 2340
	GATGACAAGC CCACCATCAT CCGCGTGGGG AAGAGCTCC TCACTCTBAC CCCCTCTGCTA 2400
	GGGCTCACCT GGGGCTTTGG AATAGGAACA ATAGTGGACA GCCAGAATCT GGCTTGGCAT 2460
20	GGTATTTTTG CTTACTCAA TGCATTCCAG GGATTTTATA TCTTATGCTT TGGAACTCTC 2520
	TTGGACAGTA AGCTGGGACA ACTTCTGTTT AACAAATTGT CTGCCITTAAG TTCTTGGAG 2580
	CAACACAGAA AGCAAAACCT ATCAGATTTA TCTGCCAAAC CCAAAATCTC AAAGCTTTTC 2640
	AACCCACTGC AAAACAAAGG CCATTATGCA TTTCTCATA CTGGAGATTG CTCGGACAAC 2700
	ATCATGCTAA CTCAGTTTGT CTCAAATGAA TAA

Seq ID NO: 175 Protein sequence
Protein Accession #: Eos sequence

30	1 MKVGVLWLIS FFTFTDGHGG FLGNDGIKT KKELVNKKK HLGPFVEEYQL LLQVTYRDSK 60
	21 EKHDLENFLK LKPKPLLSH GLIRIIRAKA TTDNLSLNGV LQCTCEDSYT WFPSCILDFQ 120
	31 NCYLHTAGAL PSCECHLNNL SQSVNFCERT KIWGTFKINE RPTNDLNSG SAIYSKYANG 180
	41 IELQLKAYE RIQGFESQVQ TQFRNGSIVA GYEVVGSSSA SELLSAIEBV AEKAKTALHK 240
	51 LFPLEDGSPR VPKKAQCNDI VFGFGSKDDE YTLPCSSGYR GMITAKCESS GWQVIRETCV 300
35	61 LSLLESLNKN PSMIVGNATE AAVSSFVQNL SVIIRQNPST TVGNLASVVS ILNIISSLSL 360
	71 ASHFVRVNSFT MEDVSIADN ILNSASVINW TVLLREKCYA SSRLELTLEN ISTLVPTAL 420
	81 PLNFSRKFDI WGIPIVKNQK LKRGYSYQIK MCPQNTSIP RGVKLIGSDQ FQRSLPETII 480
	91 SMLSLTLGNI LEVSKNGNAQ VNGPVIISTVI QNYISINEVFL FFSKLBSNLS QPHCVFWDPS 540
40	101 HLQNDNDGCH LVNETQDIVT CQCHLTSFSS ILMSPFVPSI IFFVVKWITY VGLGISIGSL 600
	111 ILGLIIEAIF WKQIKKSQTS RTRRICMVNI ALSLLIADVN FIVGATVDTT VNPBGVCTAA 660
	121 VPFTHFFYLS LFFMMLMGI LLAYRIILVF HEMAGRLMMA VGFCLGYGCF LIISVITIAV 720
	131 TQPSNTYKRR DVCLNWSNG SKRILAFVVP ALAIVAVNVP VVLLVLTKLN RPTVGERLSR 780
	141 DDKATIIIRVG KSLILITPLL GLTWGFGIGT IVDSONLANE VIFALLNAPQ GFFILCFGIL 840
	151 LDKLEQLLP NKLSALSSEK QTEKQSSDL SAKPKFSKPF NPLQMKHYA FSHTGDSDDN 900
45	161 IMLTQFVSNE

Seq ID NO: 176 DNA sequence
Nucleic Acid Accession #: AB035089.1
Coding sequence: 9845..10219

50	1 GGCATGCAG CCATCGGGGA AAATCCATAG TGCAGATAAA GCAAGGAGGA AGAAGAAGGA 60
	11 CAGTTCTAGT AAAAGGAGGA ACATCAATAT AGGATGTTTC TTAGCAATAG AAAAGAAGG 120
55	21 CCAAGAGGAA TTAGGGAGAG AGTTATAGGA GATCAGCAAG GGGACAGGGT TAGATTGGT 180
	31 TTGGTTTGRA AGCATACAGT AAATATGATG TCTGTCCCTG GCAGTGTGG CAGAGTAGGA 240
	41 AGGAGGAAGG GAGGCAAGAG ATAATATCAT TTTCTCTGTG CTCCAACTGT ACTTACATAT 300
	51 GAGACTATTT CCTCTCTGCG TTTTCAAACC TTAGTGGAGT TGTITTCCTT CATGAAAACC 360
60	61 AAGAAAGGAA AGCTAGTATG TCTTGTCTG AGGTGTGTCA ATGTATACAT ATCTATATCT 420
	71 GTAGACAGAA TCTTGGGAAA TACAGTAATT GACATATATT CTGTATTTG ATGCTTGAAA 480
	81 AATCTCTCTC ACTAACCACT TTCCCTATAG ATTGCCACAA GCACATAATA AGAAACAATA 540
	91 AATAAAATGT TCTCTGACT TTGTACTTA ACAATGCTGA GAAACCTTA CAGCCTTCAT 600
	101 AAGGAAGTGA GGTTCAGGAA AATCTAGGAG ATATTCTTA ACCAATCTAT AAAGGCATTA 660
65	111 GTAATGACAG GATATTCTCT GAAAGGTGAA TTTCCCATG AGGATTTGTT TTTAATTTCT 720
	121 GGATTCCTGG AGGCAATGAA GTTGGTGTAT GTTTATGAAA TATCAAGAGA CATAGTTGG 780
	131 CAGGTGTTCA TATGCAAAAA CTCTTGGGAA TTTCTGAGTT CTCTGTGGCA ATATATGACA 840
	141 TCAGGATATG TCCAGTCTCA CACACCAGGA FATGTCTCTT CTAGCCTGTC TATCATATGC 900
	151 TAGGAGAACT ATTTAGGAAC AGAAAAAAT GCTGAAATG ATTTCTCAT TGAACCTATC 960
70	161 CAAGCTTTCT CTAATTTTAA GCAACTCTCT GGTATTTTC AGTTAGTACC TTTCTTAA 1020
	171 TTCAACCTTC AGGCAAACT TCCGTGCCCT AGACGTTTAG CCATAGTCTG AATATCTCTT 1080
75	181 CCATAGATTG GTCCCTGTA ACCCGGTTT GTCTCAGCT GTTATCTCT TTTTCTCTTC 1140
	191 CCTCCATTCC CAGGATGAGC TTGTGCTTC TGTCTATGA GACATTAGAT TCTTTCTCTT 1200
	201 TGGTACCGGA GTAAATCCAT CCTACTCCAA TAGAGGAAGG TCCATTTTGG TCTTATAGCG 1260
	211 CTGGATGCGA ACTCAGCTGA GAGGACCAAT ATTCATTTT GGAATTTCTT ATCTCAGATA 1320
80	221 TTTCCTCTTC TTTCCTTTTC TTCTATCTTT GGATTTTAG TCCATCAAAG CCCCATAGT 1380
	231 CTATTCCCGG ACTTCAATCA GGGAACTTAT ACCTCTTAAA CTCATTGAGA GACTCAAAAC 1440
	241 ATATATATTG ATACAGAGCA CCTAAGAAGA GCATGTCTTG GGGGTGTAGG AAACAGGCG 1500
	251 GTGAGAAATT TCCAGATTGG AAACACAGCT TCTTCTCTCC CATCCAGCCC CTACTTTCAG 1560
	261 CCTATGTGTT TCTGGCACCT TGTGTAGAT AAATCTCCCT TGACTTTGTG ATGTGCTGAG 1620
	271 AAAACAACCT CAGGCTTGGT GTTAAAAAGG GCCCATGACA ATACCAAGTG TTGGGGAGAA 1680
	281 TGTGGAGAAA TCAGAACTCT ATTCACGGTC GGTGGAATG CACACTGTG CAGAATTCTA 1740
	291 TGGAGAAGAG TCTGGCATTT CCTCAAATG TTAACCTGGA TTTACCATAT GACCCAGCGA 1800
	301 TTTCATTCAAT AGGTTTATAC TCAAAAGAAA TGAAGAAATA TGCCATGCAA AAAAATGTAC 1860
	311 ATGAAAGGTC ACAACATCAT TATTCATAAT AGTAAAGGAA TGGAAACAAC ACAAAATGTC 1920
	321 ATCAACTTAT GATTAAAGAA AATCTGGTCT ATTCATAGAA TGGAAATATA TTGACACACA 1980

	AAAAGGAATG	ATGTACTGAT	CCATGCAATG	ATGTGGACAA	ACCATGAAAA	TAACACTAGA	2040
	TTAAAGAAGC	CAGTCACAAA	AGGACTTACT	GTATGATTCC	ATTTACCTGA	AAATGTTTGA	2100
	ATAGGCAAA	CCATAGAAAC	AGGAGGTAGA	TTCTTGGT	CCAGGCTCTC	CAGGAAGGGA	2160
5	AGAATGAAGT	ACAAGATTTC	TTTGGAGGT	AGTGAAATTG	TGTGGGAATG	AGATCATGAT	2220
	GATGATAGCA	CAACTTTGTG	AAATATAATA	AATCATTGAA	TGTACAGTT	GAATTTATGG	2280
	TATATAAATT	ATATGTTAAT	AAAAAGGGGG	TCCACAAAAC	AAACAGCCCC	CCACTCTGGT	2340
	TGTCAGGGAG	ATATTGGATT	AAATGGCCTT	GGACAACAC	CCCTCTCCCT	GGCCACAGAC	2400
	ATTTCTTACA	TTACAAGATA	TTCCAGGGGA	AACACTGGAA	TGAGTCTGAA	GCCAGGTGCT	2460
10	AAACAGAAGG	ACCATTTGAG	AATGTTGTGA	TCCTGACAGG	TCAAGCAATT	TATTTTTTCGG	2520
	CTTCATTTT	AAATGTAAAA	TTAGAAAGCT	GCCATTTAAA	ATGGCCCGTC	TGTTTCAATT	2580
	CTCTCTCTCA	GTGTCAAGCT	GTTAACTCAA	TGTGTTAGTC	TGTTTTCATG	CTGCTGATAA	2640
	AAACATACTT	GAGACTGGCA	AGAAAAAGAG	GTTTAAITGG	GCTTAGAGTT	CCACGTGATT	2700
	GGGGAGGCC	CAGATTCACA	GTAGGAGGCA	AAAGTTATTC	TTACATGGTG	GCTGCAAGAG	2760
15	AAGATGAGGA	AGAAGCAAAA	GAAGAAACCC	CTGATAAAC	CATCGGATCT	CCTGAGGCTT	2820
	ATTAACATAT	ATGAGAATAG	CACAGAAAG	ACCGGCCCCC	ATGATTCAT	TACCTCTACC	2880
	TGGTTCCTCT	CTATCTAATG	TGGAATTTCT	GATAGATACA	ATTCAGTTG	AGATTTGGGT	2940
	GGGAACACAG	CCAAACCATA	TCACCTAGCA	AGGCAGATAA	CTTCTCACT	GAGCCTATBC	3000
	AACAGAAAAC	CATCTGGGAT	GGTGTAAAG	GGCACAGGAA	GTGACTGGTA	GGATCACTGC	3060
20	CAGAGCTGAG	CACCTAGGAG	AAGGCAATAG	AATCTATTTC	TCCATAGTAT	GCTATAAGAT	3120
	ACTGAGTAC	ACTTCTCTAC	TATCTCTTTG	GACTTAGAAT	TAGCACTACA	TTCTTGTGTA	3180
	TACAGAAAAA	TTACTAGGGA	AATTCATAG	ATGACAAAAA	CTTTCAGAAC	TGAAAAACAG	3240
	GAATGTAAAG	CTTTTATGTT	CTTGGTATTT	CGAAGTATGC	CTAAAAGACA	ATGCAAAATC	3300
	CAGAAAAAGA	ATGGTGGGGT	TTTTGTTTGT	TGGTTTTTGT	TTTTGTTTTA	CAGCTGGAGT	3360
25	AGAAATACAA	GGGATGGAGT	TGAACAAAT	GAGAGGAAAT	TGGAATTTCTA	AACTTATTCT	3420
	CATTGGCAAT	AGAAAGGCAC	CTACATGTAT	TTCACATGAG	CCGGTGACTG	CTGACTTTGA	3480
	TTCTTATTTT	TTCCCTATAG	ATTAAAAAGG	AGGTACAAATG	GTAGAACTGT	AATCCGTGTC	3540
	TTTGTCTATA	ATTTTCTAT	TCATAAAGGT	GAGTGTATAG	CCGCTTGTGA	AATCTGAAGT	3600
	TGAGTAACTT	CABAATCTAA	CCACAGAGGG	AAAGGCAGCA	AGAGGAGAGG	CATAAATTTA	3660
30	GGATCTCACC	CTTCTATTCCA	CAGACACACA	CAGCCTCTCT	GCCCACTCTC	GCTTCTCTTA	3720
	GGACACAGG	TAAAGGCTTC	AAAGCCTCTCC	AGCTTAATAA	CATGAATTAT	TTTTGAGAAAT	3780
	AATATGATG	CTGTGTCTCA	TATCATGCAT	CTCCTGCAT	CTGTCTGATT	ATATTTTACT	3840
	TATTTCTGCA	GAGCAAAAT	AAATACCTA	TTTCATCTGA	TTTGTCTTTT	ATCTAAATG	3900
	CTTAGTTTCA	AGTAAACCAA	GGCATTTTTA	GGAACACAGA	GAGAGAGTGC	CTTGCAGCCA	3960
35	GAGAGTCTTG	AAGGAGATGT	CAGGGACGCA	TCTTAACAGC	TGGTGGATG	TGATCCACAG	4020
	AGGCTCTCTG	TTAGCATTTCA	TGTAAAGGCC	ATCCTACCTA	GCTCTAGTGT	AACCAGCAAT	4080
	GAAAGAAAGA	TAAAGAGGGT	CBATTACTTA	TTTACAATAG	TCTTFAAAAA	CSTAGTTTTG	4140
	TAGCCTTCT	AATTAAGACA	TTAATATATT	TAATATATGC	ACATTGTAGA	AAGATTGAAG	4200
	CGTTAAATAT	AAGAGAAAAA	CTTTAAATGT	CAAAATCTCA	CAACCCAGAT	ATATCAITTC	4260
40	TTTAAGAAAA	TTGFACTACA	AAATACCATT	CCATTTATTA	AAGTCATTCT	GACAGGAATC	4320
	TGATGCTTTT	CCAGGAGTTC	CAGATCACAT	CGAGTTCAAC	ATGAATTAC	TCNGTGAAGC	4380
	CAACACCAAG	TTGCTTTCTG	ATCTGTTTCA	ACAGTTTACA	AAATCAAAAG	AGAACAACAT	4440
	CTTCTATTCC	CCTATCAGCA	TCACATCAGC	ATTAGGGATG	GTCTCTTAG	GAGCCAAAGA	4500
	CAACACTGCA	CAACAAATTA	GCAAGGTAGC	TATCAGCATC	ATTACGTTGT	CCTGTTGCMG	4560
45	TTTTTCTCTG	GTTCCGTGCG	CTAGCAAGCA	GATGGTAATA	GATGTGGTGG	TCIGATGGGT	4620
	AGCACAGGGG	GCTGTGCGAG	AAATCCCATTA	ACTGTGAGAC	CACCTGACTA	AACAGATCTT	4680
	TGAGTAAGT	TTTTTCTGTC	CCGCTTCATG	TCTCTTCCAG	GTCTCTTACT	TGATCAAGT	4740
	CACAGAGAAC	ACCACAGAAA	AAGCTGCAAC	ATATCATGTG	AGTCACAGAG	CACCTCTGAT	4800
	CAGCTTTAGA	TCCCTGAAACA	GGTCATAGTT	TAAACCTGGA	ACTTCACAAA	AACTAAGAAA	4860
50	AGGCCAGTTT	TAGGGAAAT	CTTGGACACA	AAGATTGAGA	CATACAGAGT	GGGTGGGCAT	4920
	TTCAATGGCAC	ATAATTAATA	TTCCCTCATTT	CTGCGTTACT	AAAAGACAGT	CAGCACTGTA	4980
	CTTCAGAGCA	TAGGCTTGGA	TCAGGATAGG	CTGGGTTTCA	ACTCCAGCTT	TGCTCTTCAC	5040
	AAATGATGAA	TAAAGCAGG	ACACAACCTC	TOGGAGTCCC	AGTGACCTCA	TCCCAGAAAA	5100
	CTAAGGGTAA	GAAAAATCT	GACTCAATAC	ATGCAAAATG	ATGCAAAATG	TTACACAGT	5160
55	GCCTTGCCCA	TAAAGTCTAT	AATTAATGTT	ATTATTATTA	TAAAGTAGCT	ATAATTATAC	5220
	TATCTAATAT	AATGTGAAAA	TAAATTAATT	TTCAATGAGT	CATTAAATGAG	ATTCAGAGGA	5280
	ATAGACACAA	GCTCAAGTAT	ATTTTGGAAA	ATGATTGCTA	TGGAATATAT	TGGTTTAGAG	5340
	CCTTAATAT	GCAAAATGCT	TGCTGGGAAG	GTAGAAAGTT	CTAGATTATA	ACAGGCTTAG	5400
	GTTCAAACT	TGGCACTTCT	AAITTAATGTC	TCTATAAACA	GGGTTTTTTT	CCCCATTTCT	5460
60	TGAGCTTTCT	TGTGTTTATC	TGAATTGATC	TAAAGACTTA	GAGTTACCCA	TGTAAAGTCC	5520
	TTAGCCATGG	ACCTGGCATA	CACCTCTCTT	ACGTGCAGAG	AAAGACCATC	ATGAGGAAAG	5580
	AGCCACAGAT	CAGTCAATGT	GTCTTACAG	ATAATAGCAC	CAACAGTAT	AACAGGCTT	5640
	CCTGGCATAA	TCTATTATAA	ATATCCAACC	TTCAACATAC	TGTTATCTCT	GATGACTGTT	5700
	AGAAGTGAAA	TATGGTCTCT	GCCCATAGG	AGCTGAGAGT	TTAACTGGGA	AGCTAAACCT	5760
65	AAOCTTTAA	ACCAACAAGG	AGAAATCTA	CTGGTAGACA	GCGCTGCATC	TTTAGTTTCA	5820
	AAGAGAAAG	ATTGCAAGT	GTAGAGCAA	GAAGAAATTT	CTGGAAGAG	TCAAAATATA	5880
	GGTGGATTTT	GAAGGGTATT	TGAGGTGAAA	TACACCAATT	ATCAGGGAAT	AACATCAAG	5940
	GTCTCTCAAT	AGACTACCA	CAITTAGGGA	CTGATCTAAC	AGACTTAGCA	TGGGTTTAGT	6000
	ATTTACATG	ATACAGCAAT	TGAATGATCT	CTTTTTTGA	TGTTTGAAG	TGATAGGTC	6060
70	AGGAAATGTT	CATCAACAGT	TTCAAAAGCT	TCGACTGAA	TTCAACAAAT	CCACTGATGC	6120
	ATATGAGCT	AGATCGCCA	ACAAGCTCTT	CGGAGAAAG	ACGTATCAAT	TTTACAGGT	6180
	AAITTCACCT	GGCTACCCA	CATTTCATT	GCACTCTGAT	GTCTGTGTCT	CTGAGTGGCC	6240
	AAATGGAAGA	AAGCAAGGCA	GATGAGCTTG	GCCGACCCAG	GTGGAGAGCA	TTTACTCAGA	6300
	GTGCTTAGC	TCCATTTCCA	CAACTCTCCC	CCACTGGAGT	GTCCCAAGCC	CCACGATAC	6360
75	ATCACTGAG	TGTGGATTTA	GGGATAATCT	TGTGATAAAA	GAGGAGGTTG	TGTAATAGAG	6420
	TGAGTAAAG	TAAATAGTAA	TAGATACCA	TGATAAACT	GGCACTGACT	CAGTCACATA	6480
	CGATACATG	TGGTGGGAAA	TGTATGACTA	ATGGGATATT	ATTGGAAATG	GCAGGCTTGG	6540
	GTGAGTTCTT	GAGAAATAGT	GAGGAAGTAC	CAGGAAATAT	TGAATGCACA	GGATGAAAGA	6600
	CAAAAACAAA	GATCAGAAAC	ATCATGGTTA	AAATTAATCG	AGAGAGTCT	GAGAGCAAT	6660
80	GAATCTCTT	CAGGGAAGCC	TGCTCTGCAG	TTTGCAAAAC	ACAGCCTCTT	CTGCTTCTGC	6720
	CTTTTGCCAA	GATGATATTG	ACCTTCAGTG	ACCTCTTCT	TGTGCCAGCC	CACATTCCTC	6780
	TTTGTCAATG	CCTACATGAC	ACCTGTATAA	AAATATCCAT	GGACAGGAGA	TACTGCTATC	6840
	ATTCAGGCTC	TGATTTAGC	TTACTGTTGT	TACAAATAAG	TAAATTTGGT	AAATATATAG	6900
	TACATAAATT	ACTCTTAATT	CCATCTCTT	CTTCAATATC	TCAAGGAAT	ATTTAGATGC	6960
	CATCAAGAAA	TTTTTACAGA	CCAGTGTGGA	ATCTACTGAT	TTTGCAATG	CTCCAGAGA	7020

AAGTCGAAAG AAGATTAACT OCTGGGTGGA AAGTCAAAAG AATGGTAGGA GAGCCACCCA 7080
 TTATAGAAAC ACCTTTGAGA AACCTATGCC AGTGAGCCTT GTGCTTGACA CTGCATGGGG 7140
 GAACAGGTGT GGGGATTGAG ATGGGTTTGC AGGGAGGGCT GAAGAGGGCA CTCCAGATGA 7200
 AGGATTGTCT CAAATGAATA TGAAGAGAGC CTAGGGGAGC CAAGGAGGAA ATCACAGGAA 7260
 GCCAATTAGA TGGAAACACA TCTGGAGAAAT TATTTGCTTA TGGCCCTGCA TGACAATAGC 7320
 TTTGTGGATC CCTGTCTCC GCTCAGACCT ATTTTGAGAT CATATCCTTT ACTTTAAATC 7380
 AGACTCAAT TTTTATGATG AATATTTAAT AGAAACATT AGAAAGCGTC TCTGTCTCC 7440
 TTTACTAAT GGGAAACAG CAGCTCTCTG GTAAATCACC CTTTGTCTC TGAGCTGGAG 7500
 CTGCCCTGGT CACATCTGTA GCCAATGTGT TCTGCAGGGA TTATCAGAGC TCTCTTCCCC 7560
 ATCAAGGGCA AAGAGCTTGA CAAAGTCTCC ATTCTACAGA CATCTTTCTT ACCTCCACCC 7620
 TCTCATTA CAAGCAAACT ACAGCACTC AACATGAGAG TGAATAGGAA GATACCCCCG 7680
 GAAGTAGAT CTGACAGCAC AGGACATGCG TTTTCATATTA CAGAGCTCAA GTCACTCATC 7740
 CTAAATGCA ATCAGGCGCT CCTTCTCTG ANTGGGACC CCGTAGTTAA AAAAAATAA 7800
 AAGTAGAAG AGGAGGGAGG GAGAAAGGAA AGACACATGT TGGACAGTA GACAAAATCA 7860
 GTTATCAGT TTTTCAATC AGATGATTGG AGACATTCA ACACAGAGAA CGTGAACCTC 7920
 TTTCTCATCA CAAGAATGAA TGTCTCCATC AAGGGTAACT TTATACGACT GGAGCCTTGA 7980
 AGAAAGCTGC ATTTGGGTAA CCACCTGGTCA GTGAGTCTAA CAATTCAAAG ATCAAAGTCA 8040
 GTGAGTCTCA AGCAGGGATT TGGGTCAATA ATTAACGATC AGTCACGAAC ATTTGCAAAG 8100
 CATCTTCCAG ACAAGCCATT TGTAGCTTGT GTAAAAGACT CTTTATTCT TCCCTTGCA 8160
 GAAAAAATA AAAACCTATT TCTGTATGGG ACTATTGGCA ATGATACGAC ACTGGTTCTT 8220
 GTGAACGCA TCTATTTCAA AGGGCAGTGG GAGAATAAAT TTAATAAAGA AAACACTAAA 8280
 GAGGAAAAAT TTTGGGTCAA CAAGGTATTG TCTATATTTT ATTTATATAG TGAATATGT 8340
 TAATACATGG AATGTTAAAC ATTTCTGATG GAATGTAA CA TGAATAGTAA AAAATAAAAA 8400
 TGTTCATGT CTTTATTCT GTTGTCTTAC TCTTATAACT TTATTAGTT AGGAATACCT 8460
 GAAAACTAT TGTCTTAAC TCAATGAAAT OCTGGGTTAT TTCTTAGAAG AGAAGGATG 8520
 TGTGTCTATC TCAATATAT TATCTTTT GTCTGTGTT TCACTGTGTA TTTGTGGAC 8580
 ACATTGATT ATTTGCTGAT ACATACAAAT CTGTACAGAT GATGAGGCAA TACAATCTCT 8640
 TTAATTTGCT CTGCTGGAG GATGTACAGG CCAAGGCTCT GGAATACCA TACAAGGCA 8700
 AAGATCAAG CAGATTTGTG CTGCTGCCAA ATGAATCGA TGGTCTGCG AGGTAAGAA 8760
 CTGCACTCA CAATCTCTCC TTCTACTGCC GGACATTTT CCAAGATAC CAAGTTTAAA 8820
 CAAGGIAAAA GCTTATGACC GAGTTGCCCT AAAATGATGA AAAATCTCAA ATGAGGAAATG 8880
 ATGACTCAC TTTGGGTCAA AAATATTGA GCATAGGGCC TGACACAAAC TGAAGCTTA 8940
 GTTTTGTGT GTTGTGTGT TTTTATTAT ATTATTATA TACTTTAAGC TTAGGGGTAC 9000
 ATGTGCAAAA TGTGAGGT AGTTACATAT GTATACATGT GCCATGCTGG TGTGCTGCAC 9060
 CCATTAATC ATCAATTAGC GTTAGGTATA TCTCTAATG CTATCCCTCC CCCCCTCCC 9120
 CACCCCAAA CAGTCTCAG AGTGTGATGT TACCTTCTG TGTCCAAGTG TTTCTATTGT 9180
 TCAATCTCA TCTATGATT AATTCATCT ATGGCTTAGT TAATGATTAA TTTATTAGAG 9240
 TTACATGCT TGGATATCAA TTTGATGATA TTATTATGCA GCAATTTAA CTTGACTGGG 9300
 AGAATATAT ACCAATGTGA GGAAGTTTA CAATAGGCC GAGTAGAAAA GGAATACAA 9360
 ATTTAGGAAT TTAGGGAAAT ACAATTAAT AATTGCAATG TGTACTAAAT AATGTATACA 9420
 GAAAAATATG ATGAGCCTAT TAAAAATGA CACATGTAGT AGGCTGTGG CACAAGAAAT 9480
 AGTGATACAT TGTGTGATT GTGTACAAAA TAATGTAATC ATATTTTACA TGTGTACTT 9540
 ACAGTTGAT ACATACATAT GTACACATAT ACATATAGT AAAACATGA TTTCTTTTT 9600
 ACATACATGT ATATACATAT ACACATATA CCAATGTAT TTATATTTT AGGACTATA 9660
 TTTTACCTAT TAGAATAATA ATGTCTATTA AAGTGAACCT TCTGTATTTC ACATTTATTG 9720
 CCAAAATAAC GAATCTCCAC ATAGTCAAT CATTTGTAAG GTGTATTAGA GATCGACAGT 9780
 TAGTATATC ATTTCTTTT TTCCATTTGT ATAGCTTGAA GAGAACTCA CTGCTGAGAA 9840
 ATGTATGGA TGGACAGTT TGCAGATAT GAGAGAGACA TGTGTGATT TACACTACC 9900
 TCGGTTCAA ATGGAAGAGA GCTATGACCT CAAGGACAG TTGAGAACCA TGGGAATGGT 9960
 GAATATCTT AATGGGGATG CAGACCTCTC AGGCATGACC TGGAGCCAG GTCTCTCAGT 10020
 ATCTAAAGTC CTACACAGG CCTTGTGGA GTTCACTGAG GAGGAGTGG AAGCTGCAGC 10080
 TGGCACCGCT GTAGTAGTAG TCGAATTATC ATCTCCTTCA ACTAATGAAG AGTTCTGTTG 10140
 TATCACCCCT TTTCTATTCT TCATAAGGCA AAATAAGACC AACAGCATCC TCTCTATGG 10200
 CAGATTCTCA TCCCATAGA TGCATTAGT CTGTCACTCC ATTTAGAAAA TGTTCACCTA 10260
 GAGGTGTTCT GTTAACTGA TTGCTGGCAA CAACAGATT TCTTGGCTCA TATTTCTTTT 10320
 CTATCTCATC TTGATGATGA TAGTCATCAT CAAGAATTA ATGATTAAAA TAGCATGCTT 10380
 TTTCTCTTT CTCTTAATA GCGCACATAT AAATGTACTT TTTCTCCAG AAAAATTTCC 10440
 CTGAGGAAA AATGTCCAG ATAAGATGAA TCATTTAATA CCGTGTCTTC TAAATTGAA 10500
 ATATAATTCT GTTCTGACC TGTTTAAAT GAACCAACC AAATCATACT TTTCTTCAA 10560
 ATTTAGCAAC CTAGAAACAC ACATTTCTT GAATTTAGGT GATACCTAAA TCTTCTTAT 10620
 GTTCTAAAT TTTGTATT C TATAAACAC ATCATCAATA AAATAATGAC ATAAAAATCA 10680
 TTTTGTCTTA CCTGTTTCT CTCTGGAAAG GGCAGTGT CAGTTACACA TAGGAAGAT 10740
 AATTTAGAGA TATATTATC ATATATAAG GAAATTAATA AACAGAGTAG TTCATGATGA 10800
 GCCTGGAGTA GAAGGCATAT CCGAAGACAG GAGGAGCCTT GTAAACCA CA TAGGACTTC 10860
 CTATTTTATG CTAAAGGGAT AAGAACTCA TTACAGGCTT TGATGGTTGT TTGTCAAGA 10920
 GGGGCATAAA ATTATCATAT CCACATCTAG AAAATACATC TCTGGCTACG CTGATATCAA 10980
 TGGATGCGAG GAAAGAACAG TGTGGTTACC ATATATAAAT TAGGAAATCA TTAGATATT 11040
 GGGAGTGGAA ATGAGAGAA AGAAAGAGCC TGGGGGAATT ATTTAGGAAA TAATAGTTAC 11100
 AGAAAGCAT CTAAGTTGCT GACCTATCTG ACTGGATGGA TGGAGAATA TCTTGTCTT 11160
 GAGAGAAAAA AAGACTTTGG GTTTAAATTT GTACTTGATG AATTAGGTA CTTTAAATAT 11220
 TCAATGGAT TTTCCCTGCA GGCCTTGAA GATATTAGTC TAAATCTCAG AAACAGATA 11280
 TGTACTGAAG CTCTAAATTT GTGATATTCA ATATAAATAC TTTAGAGTCA TGGGATAAA 11340
 TATGGTAGTT GTAGCTAAAA GCAAAAATAA GATACTAGG AGAAAGGATA AAGTTAGAG 11400
 AAAGAAGAA CTAGAATTGA CCTTGAAGTA TATCAGCATG TGTAAAGATC AGGAATTGAT 11460
 CATTTTATT TTCCAGAAAG TAGCTTTCT TAGGGTCCA TATTTACTCC CATAGATTCT 11520
 TCCC

Seq ID NO: 177 Protein sequence
 Protein Accession #: BAB21525.1

1 11 21 31 41 51
 MNSLSEANIK FMDLFQQR SKENNIFYS PISITSALGM VLLGAKDNTA QQISKVLHFD 60
 QVTENITEKA ATYHVDKSGN VHQPOKRLIT EPNKSTDAYE LKIANKLFGF KTYQLQEXL 120

DAIKKFYQTS VESTDFANAP EESRKIKNSW VESQTNKIK NLFPDGTIGN DTTLLVLNAI 180
 YFKGQWENKF KKEKTKBEKF WPNKNTYKSV OMNRQYNSFN FALLEDDVQAK VLEIPYKSKD 240
 LSHIVILLPNE IDGLQKLEEK LTAEKLMWENT SLQNMRETCV DLHLPRFKME ESYDLKDTLR 300
 TMGMVNIWNG DADLSGMTWS HGLSVSKVLH KAPVEVTSBG VEAATAAVV VVELSSPSTN 360
 EEPCCNHPFL PFIRQNKINS ILFYGRFSSP

Seq ID NO: 178 DNA sequence
 Nucleic Acid Accession #: NM_001910.1
 Coding sequence: 50..1240

1 11 21 31 41 51
 | | | | |
 GGAGAGAAGA AAGGAGGGGG CAAGGGAGAA GCTGCTGGTC GGACTCACA TGAACACGCT 60
 CCTCTTTTIG CTGCTGGTGC TCCTGGAGCT GGGAGAGGCC CAAGGATCCC TTCACAGGGT 120
 GCCCTTCAGG AGGCATCCGT CCTCAAGAA GAAGCTGGGG GCACGGAGCC AGCTCTCTGA 180
 GTTCTGGAAA TCCCATTAAT TGGACATGAT CCACTTCACC GAGTCTGCT CAATGGACCA 240
 GAGTGGCAAG GAACCCCTCA TCAACTACTT GGATATGGAA TACTTCGGCA CTATCTCCAT 300
 TGGCTCCCCA CCACAGAATC TCACTGTCTT CTTCGACACT GGCTCTCTCA ACCTCTGGGT 360
 CCGCTCTGTG TACTGCACTA GCGCAGCCCTG CAAGACGCAC AGCAGGTTC AGCCTTCCCA 420
 GTCCAGCACA TACAGCCAGC CAGGTCAATC TTTCTCCATT CAGTATGGAA CCGGGAGCTT 480
 GTCCGGGATC ATTGGAGCCG ACCAAGTCYC TGTGGAGGGA CTAACCCGCG TTGGCCAGCA 540
 GTTTGGAGAA AGTGTCAAG AGCCAGGCCA GACCTTTGTG GATGCAGAGT TTGATGGAAT 600
 TCTGGGCTCG GGATACCCCT CCTTGGCTGT GGGAGGAGTG ACTCCAGTAT TTGACACAT 660
 GATGGCTCAG AACCTGGTGG ACTTGGCTAT GTTTCTGTCT TACATGAGCA GTAACCCAGA 720
 AGGTGGTGGG GGGAGCGAGC TGATTTTGG AGGCTAGGAC CACTCCCAT TCTCTGGGAG 780
 CCTGAATTGG GTCCAGTCA CCAAGCAGC TTAAGTGGAG ATTGCACTGG ATAACATCCA 840
 GGTGGGAGGC ACTGTATGAT TCTGCTCCGA GGGCTGCCAG GCCATTGTGG ACACAGGGAC 900
 TTCCCTCATC ACTGGCCCTT CCGACAAGAT TAAGCAGCTG CAAACGCCCA TTGGGGCAGC 960
 CCGCTGGAT GGAGATATG CTGTGGAGTG TGCCAACCTT AACCTCATGC CGGATGTAC 1020
 CTTCAACATT AACCGAGTCC OCTATACCT CAGCCCAACT GCCTACACCC TACTGGACTT 1080
 CGTGGATGGA ATGCACTTCT GCAGCAGTGG CTTTCAAGGA CTTCAGATCC ACCCTCCAGC 1140
 TGGGCCCTCT TGGATCTGG GGGATGTCTT CATTGACAG TTTTACTCAG TCTTGGACG 1200
 TGGGAATAAC CGTGTGGGAC TGGCCCGAGC AGTCCCTAA GGAGGGCGCT TGTGTCTGTG 1260
 CCTGGCTGTC TGACAGACT TGAATATGTT AGGCTGGGCG ATTCCTTACA CCTACAAAA 1320
 GTTATTTTCC AGAGATGTA GCTGTCTTCA GGGTTCGAAC TTGAATTAG ACCAAACAGA 1380
 ACATGAGAAAT ACACACACAC ACACACATAT ACACACACAC ACATTCACA CATACACACC 1440
 ACTCCACCCA CCGTCTGAT GGAGGAATTA CGTTATACAT TCATATTTTG TATTGATTTT 1500
 TGATATGAA AATCAAAAAA TTTCACTTT GATTATGAAA ATCTCCAAAC ATATGCACAA 1560
 GCAGAGATCA TGGTATAATA AATCCCTTGG CAATCCACT CAGCCCTGAC AACCCATCCA 1620
 CACACGGCCA GGCCTGTTTA TCTACACTGC TGCCCACTCC TCTCTCCAGC TCCCATGCT 1680
 GTACCTGGAT CATCTGTAAG CAAATTCGGA GCATTACATC ATTTGTGCA TAAATATTTC 1740
 TAACATCCTT AATATACAA TCGGAATTCA AGCACTCTCC ATGTGCCAC AAATGTTTGG 1800
 CTGTTTTGTG AGTGGATTTG TTTGTATTAG GATTCAAGCA AGGCCCATAT ATTGCATTTA 1860
 TTTGAAATGT CTGTAAGTCT CTTTCCATCT ACAGAGTTTA GCACATTTGA ACCTTCTGCG 1920
 TTGAAATCCC GAGGTGTCTT TGCATATGTT TCTCTGAACT TATCTTCTCT ATAAATGGT 1980
 AGTATGATCT GAGAGTCTGA TTTTGTGGCA AAAATATCTC CTAGGTGGTG CTGGGTACTT 2040
 CTTGTGATC CTGTCTCAGA GGCAGATAAT GCTGGTGCTT CTCTATTGGT AATGTTAAGA 2100
 CTGCTGGGTG GGTITGGAGT TCTTGGCTTT AATCATTCTT TACAAAGTTC AGCATTTT

Seq ID NO: 179 Protein sequence
 Protein Accession #: NP_001901.1

1 11 21 31 41 51
 | | | | |
 MKTILLILLV LLELGEAQQS LHRVPLRRHP SLKKKLRRAS QLSEFWKSHN LDMIOFTBSC 60
 SMDQSAKEPL INYLDMEYFG TISIGSPFPQN FTVIPDTSBS NLWVPSVYCT SPACKTHSRP 120
 QPSQSTYSG TGLSLGIIGA DQSVVEGLTV VGGQFGESVT EPGQTFVDAD 180
 PDGLLGLGYP SLAVGGVTPV FDMMAQNLV DLPNFSVYMS SNPEGGAGSE LIFGGYDHS 240
 FSGSLNVPV TRQYKQIAL DNIQVGGTVM FCSBGCQAIV DGTGSLITGP SDRIKQLQNA 300
 IGAAPVDGSE AVECANLNM PDVTFTINGV PYTLSPATYT LLDVFDGMQF CSSGFGGLDI 360
 HPPAGPLWIL GDVFIRQFYS VFDRGNRRVG LAPAVP

Seq ID NO: 180 DNA sequence
 Nucleic Acid Accession #: NM_018058.1
 Coding sequence: 319..1575

1 11 21 31 41 51
 | | | | |
 TACGCGCTGC GGGACCGGCA GGGGAACGCC ATCGGGGTCA CAGCCTGCGA CATCGACGGG 60
 GACGGCCGCG AGGAGATCTA CTTCTCAAC ACCAATAATG OCTTCTCGGG GGTGGCCACG 120
 TACACCGACA AGTTGTTCA GTTCCGCAAT AACCGGTGGG AAGACATCTT GAGCGATGAG 180
 GTCAAGCTGG CCGCTGGTGT TTTGCCGGAC GCTCTGTGGC CTGTGTGGAC 240
 AGAAGAGGCT CTGGACGCTA CTCTATCTAC ATTGCCAAT ACBCCTACGG TAATGTGGGC 300
 CCGATGCCC CTCTGTAAT GGACCTGAG GCCAGTGACC TCTCCGGGG CATTCTGGCG 360
 CTCAGAGATG TGGCTGTCTA GCGTGGGCTC AGCAAAATTA CAGGGGGCGG AGGCTTCAGC 420
 GTGGGCCCCA TCTCTAGCAG CAGTGCTCTG GATATCTTCT GCGACAATGA GAATGGGCTT 480
 AACTTCTTTT TCCACACAC GGGCGATGGC ACCTTTGTGG AGCCTGCGGC CAGTCTGGT 540
 GTGACACACC CCAACAGCA TGGCGAGGT GTGCGCTGG CTGACTTCAA CCGTGTATGG 600
 AAAGTGGACA TCGTCTATGG CAACTGGAAT GGCCGCCACC GCCTCTATCT GCAAAATGAG 660
 ACCATGGGGA AGGTCCGCTT CCGGACATC GCCTCACCCA AGTCTCTCAT GCGCTCCCTT 720
 GTCCGACGCG TCACTACCC CCACTTTGAC AATGACCAAG AGCTGGAGAT CTTCTTCAAC 780
 AACATTGCTT ACCGCAAC CTCAGCCAAC CCGCTCTTCC GCGTCTATCG TAGAGAGCAC 840
 GGAGACCCCC TCATCGAGGA GCTCATCCC GCGCAAGCTT TGGAGCCTGA GGGCCGGGGC 900
 ACAGGGGTG TGGTGACCA CTTGACGGA GACGGGATGC TGGACCTCAT CTTGTCCCAT 960

5
10
15
20

```

GGAGAGTCCA TGGCTCAGCC GCTGTCCGTC TTCGGGGGCA ATCAGGGGCTT CAACACAAAC 1020
TGGCTGCGAG TGGTGCCACG CACCCGGGTT GGGGECCTTG CCAGGGGAGC TAAGGTCGTG 1080
CTCTACACCA AGAAGAGTGG GGGCCACCTG AGGATCATCG ACCGGGGCTC AGGCTACCTG 1140
TGTGAGATGG AGCCCGTGGC ACACCTTGGC CTGGGGAAAG ATGAAGCCAG CAGTGTGGAG 1200
GTGACGTGGC CAGATGGCAA GATGGTGAAG CGGAACGTGG CCAGCGGGGA GATGAACCTA 1260
GTGCTGGAGA TCCTCTACCC CCGGATGAG GACACACTTC AGGALCCAGC CCACTGGAG 1320
ACACCAATGA ATGCATCCAG TTCCATTCCG TGTGCCCTCG AGACAAGCCC GTATGTGTCA 1380
ACACCTATGG AAGCTACAGG TGCCGGACCA ACAAGAAGTG CAGTCGGGGC TACGAGCCCA 1440
ACGAGGATGG CACAGCCTGC GTGGGGACTC TCGGCCAGTC ACCGGGGCCC CGCCCAACCA 1500
CCCCACCGCC TGTGCTGCCC ACTGCCGCTG CTGCTGCCGC TGCTGGAGCT GCCACTGCTG 1560
CACCGGTCCCT CGTAGATGGA GATCTCAATC TGGGGTCGGT GGTAAAGGAG AGCTGCGAGC 1620
CCAGCTGCTG AGCAGGGGTG GGACATCAAC CAGCGGATGG AGTCCAGCAG GGGAGTGGGA 1680
AAGTGGGCTT GTGCTGCTGC CTAGACAGTA GGGATGTAAA GGCCTGGGAG CTAGACCTTC 1740
CCCAAGCCCA TCCATGCACA TTACTTAGCT AACAAATTAG GAGACTCGTA AGGCCAGGCC 1800
CTGTGCTGGC TGATCAGCAG AGACAGGGTC GCTGCCCTGA TGGCGCTTAC 1860
ATTCCAGTGG GTCTAATGAC CATATCTTAG GACACAGATG TGCCCAAGGA GTTGGTGTCA 1920
CTGCACAGTA AGTATGAGGA CTTAGTGTCT CTGAGTTCAA ATCCTGATTC AGGAACCTAC 1980
AAAGCTATGT GACCTTACAC CAGTCACTTA ACTTGTATGC CATCCATTAT CGCATCTGCA 2040
AAATGGGGAT TAAGAAATGA ATCTGGGGT TAGTGTGGAG ATTAGATTAA ATGTATGTAA 2100
GACACTTGGC ACAAACCTG GCACATAGTA AAGGCTCAAT AAAAAAAGT GCCTCTCACT 2160
GGGCTTTGTC AACCGTG

```

Seq ID NO: 181 Protein sequence
Protein Accession #: NP_060528.1

25
30
35

```

1 11 21 31 41 51
MDPEASDLBR GILALRDVAA EAGVSKYTG RGVSVGPLB SSASDIFCDN ENGNPFLFHN 60
RBDGTVDVAA ASAGVDDHQ EGRGVALADF NRDGKVDIVY GNWNGPHRLY LQMSTHGKVR 120
FRDIASPKFS MPSPVRTVIT ADFDNDQELE IFNNIAYRS SSANRLFRVI RREHGDPLIE 180
ELNPGDALEP EGRGTGGVVT DFDGDMIDL ILSHGESMAQ PLSVFRGNQG FNNWLVKVP 240
RTRVGAFARG AKVLYLTKKS GAHLRIIDGG SGYLCEMEPV AHFGLKDEA SSVVETWPDG 300
KMYSRNVASG ESNVLEILY PRDEDTLQDP APLETFMNAS SEHSCALETS FVSTPMBAT 360
GAGPTRSAVG ATSPTRMAQP AWGLSAGHRA PAPPFPLLL PLPLLLPLLE LPLLHRSS

```

Seq ID NO: 182 DNA sequence
Nucleic Acid Accession #: AJ279016
Coding sequence: 1..1962

40
45
50
55
60
65
70
75
80

```

1 11 21 31 41 51
ATGTCCAGGA TGTTACCGTT CCTGCTGCTG CTCTGGTTTC TGCCCATCAC TGAGGGGTCC 60
CAGCGGGCTG AACCCATGTT CACTGCAGTC ACCAACTCAG TTCTGCCTCC TGACTATGAC 120
AGTAATCCCA CCCAGCTCAA CTATGGTGTG GCAGTTACTG ATGTGGACCA TGATGGGGAC 180
TTTGAGATCG TGTGGCGGG GTACAATGGA OCCAACTGG TTCTGAAGTA TGACCGGGCC 240
CAGAAGCGGC TGGTGAACAT CGCGTCCGAT GAGCGCAGCT CACCTACTA CGCGCTGCGG 300
GACCAGCAGC GGAACGCCAT CGGGTCCACA GCTTGCAGCA TCGACGGGA CGGCGGGAG 360
GAGATCTACT TCTCAACAC CAATAATGCC TTCTCGGGGG TGGCCACGTA CACCGACAAG 420
TTGTTCAGT TCCGCAATAA CCGTGGGAA GACATCTCTA GCGATGAGT CAACGTGGCC 480
CGTGGTGTGG CCAGCTCTTT TGCCGAGCGC TCTGTGGGCT GTGTGGACAG AAAGGGCTCT 540
GGACGCTACT CTATCTACAT TGCCAAATAC GCTTACGGTA ATGTGGGCC TATGCCCTC 600
ATTGAATGG ACCCTGAGGC CAGTGACCTC TCCCGGGGCA TTCTGGCGCT CAGAGATGTG 660
GCTGCTGAGG CTGGGCTCAG CAAATATACA GGGGCGGAG GCGTCAGCGT GGGGCCCATC 720
CTCAGCAGCA GTGCCCTCGGA TATCTTCTGC GACAATGAGA ATGGGCTTAA CTCTCTTTT 780
CACAACCGGG CCAGCTCTTT TGTGTGAGC GCTGGGCGCA GTGCTGGTGT GGAAGGACCC 840
CACCAGCATG GCGGAGGATG CGCCTTGGCT GALTTCAAAC GTGATGGCAA AGTGGACATC 900
GTCTATGGCA ACTGGAATGG CCCCACCGC CTCATCTGCG AAATGAGCAC CCATGGGAAG 960
GTCCGCTTCC GGGACATCGC CTCACCCAG TTCTCCATGC CCTCCCTGT CGGCACGGTC 1020
ATCACCGCGC ACTTTGACAA TGACCCAGAG CTGGAGATCT TCTTCAACAA CATTCCTTAC 1080
CGCAGCTCTT CAGCCCAACG CCTCTTCCGC GTCATCCGTA GAGAGCAACG AGACCCCTTC 1140
ATCAGGAGC TCAATCCCGG CGAAGCCTTG GAGCCTGAGG GCGGGGCGAC AGGGGGTGTG 1200
GTGACCGACT TGACCGGAGA CCGGATGCTG GACCTCATCT TGTCCCATGG AGAGTCCATG 1260
GCTCAGCCGC TGTCCGTCTT CCGGGGCAAT CAGGGCTTCA ACAACAAGTG GCTGCGAGTG 1320
GTGCCACGCA CCGGTTTGG GGCCTTTGCC AGGGGAGCTA AGTCTGTGCT CTACACCAAG 1380
AAGAGTGGGG CCCACCTGAG GATCATCGAC GGGGGCTCAG GCTACCTGTG TGAGATGGAG 1440
CCCGTGGCAC ACTTTGGCCT GGGGAAGGAT GAAGCCAGCA GTGTGGAGGT GACGTGGCCA 1500
GATGGCARGA TGGTGAAGCG GAACGTGGCC AGCGGGGAGA TGAACCTCAGT GCTGGAGATC 1560
CTCTACCCCC GGGATGAGGA CACACTTCAG GACCCAGCCC CACTGGAGTG TGGCCAAGGA 1620
TTCTCCAGC AGGAATATGG CCATTGCATG GACACCAATG AATGCTATCA GTTCCCATTC 1680
GTGTGCCCTC GAGACAAGCC CGTATGTGTC AACACCTATG GAAGCTACAG GTGCCGGACC 1740
AACAGAAAGT GCAGTCGGGG CTACGAGCCC AACGAGGATG GCACAGCCTG CGTGGGGACT 1800
CTGGCCAGT CACCGGGGCC CCGCCCAACC ACCCCCAACG CTGCTGCTGC CACTGCGCT 1860
GCTGCTCCCG CTGCTGGAGC TGCCACTGCT GCACCGGTCC TGGTAGATGG AGATCTCAAT 1920
CTGGGGTGGG TGGTTAAGGA GAGCTGCGAG CCCAGCTGCT GAGCAGGGGT GGGACATGAA 1980
CCAGCGGATG GAGTCCAGCA GGGGAGTGGG AAAGTGGGCT TGTGCTGCTG CCTAGACAGT 2040
AGGGATGTAA AGGCCTGGGA GCTAGACCCCT CCCCAGGCC ATCCATGCAC ATTACTTAGC 2100
TAACTATFAG GAGACTTGT AAGGCCNAGC CTTGTGCTGG GCACATAGCT GTGATCAGAG 2160
CAGACAGGGT GCTGCGCCCTG ATGGCGCTTA CATTCAGTGG GTCTAATGTA CCAATCTTAT 2220
GGACACAGAT GTGCCAGGGG AGTGGGTGTC ACTGCACAGG AAGTATGAGG ACTTTAGTGT 2280
CTGAGTTTCA ATGCTGATT CAGGAACCTA CAAGCTATG TGACCTTACA CCACTCACT 2340
AACTTGTATG CCAATCCATT TOGCATCTGC AAAATGGGGA TTAAGAAATG AATCTTGGGG 2400
TTAGTGTGGA GATTAGATTA AATGTATGTA AGACACTTGG CACAAAACCT GGCACATAGT 2460
AAAGGCTCAA TAAAAACAG TGCCCTCTAC TGGGCTTTGT CAACACG

```

Seq ID NO: 183 Protein sequence
Protein Accession #: CAC08451

	1	11	21	31	41	51	
5							
	MSRMLPFLLL	LWFLPITEGS	QRAEPMFTAV	INSVLPPDYD	SNPTQLNYGV	AVTDVDHGD	60
	FEIVVAGYNG	PNLVLYKXDR	QKRLVNIAYD	ERSSPYYALR	DROGNAIGVT	ACDIDGGRRE	120
	EIYPLNTMNA	PSGVATYTDK	LPKFRNNRWE	DILSDEVNVA	RGVASLPAGR	SVACVDRKGS	180
	GRYSIIYANY	AYGNVGPDAL	IEMDPEASDL	SRGILALRDV	AAFAGVSKYT	GGRGVSVGPI	240
10	LSSASADIFC	DNENGNFLF	HNREGDTFVD	AAASAGVDDP	BQHGKRGVALA	DFNRDQKVDI	300
	VYGNWNGPFR	LYLQMSYTHK	VFEDIASPK	FEMPSFVRTV	ITADFNDQOE	LEIFPNNIAY	360
	RSSSANRILFR	VIREHGDPL	IEELNPGDAL	EPEGRGTGGV	VTDFDGDGML	DLILSHGESM	420
	AQPLSVFRGN	QGFNNNNLWR	VPRTRFGAFA	RGAKVVLTYK	KSGAHLRIID	GGSGYLCEME	480
	FVAHFLGLKD	EASSVEVTWP	DGKMVSRNVA	SGEMNSVLEI	LYPRDEDTLQ	DPAPLECGQG	540
15	FSQDENGECM	DTNECIQFPF	VCPRDKFVCV	NTYGSYRCRT	NKKCSRGYEP	NEDGTACVGT	600
	LQSGPGRPRT	TPTAAAAATA	AAAAAGAAATA	APVLVDGDLN	LGSVVKESCE	PSC	

Seq ID NO: 184 DNA sequence
Nucleic Acid Accession #: FGENESHH
Coding sequence: 1..4794

	1	11	21	31	41	51	
25	ATGGCGTGTG	CGGGAGGACT	CCCAGCCCGT	TGCTCTGGTT	GGATGGGACT	GGGTGGGCCC	60
	AGCGGCTCCT	CCCCAGCATC	CCCTCCCAT	TCCTCCTCCA	GGTACAATGG	ACCCAACTTG	120
	GTTCCTGAAGT	ATGACCGGGC	CCAGAAGCGG	CTGGTGAACA	TCCGCGTCGA	TGAGCGCAGC	180
	TCACCTTACT	ACGCTCTGCG	GGACCGGCAG	GGGAACGCCA	TCCGGGTCCAC	AGCTTGCAGC	240
	ATCGACCGGG	ACGGCCCGGA	GGAGATCTAC	TTCTCTCAACA	CCAATAATGC	CTTCTCGGGC	300
	CACAGCAGCT	CAGCGCAGGT	CCCTCTCTGG	CTCCACAGAA	ACAGGCGCTGT	GCTGAAGCCT	360
30	CCACCTACAA	ACAGAAAGAG	CTCTCTGAGG	CTGGCTCCAC	TCAGCGGAAG	GGACTTTTCC	420
	TCTCTCCTGG	GTCAGGCTTC	TCCGACAGCG	AGGACGGGAG	AGAGGGTGC	GGTTCCTGTC	480
	TGTCGGGGTG	GACTGAGACC	TACCATGAA	CCAGAACCAT	TTCTCTGAG	ACCCAAATCA	540
	GGGGTGCCCA	CGTACACCGA	CAAGTTGTTT	AAGTTCGGCA	ATAACCGGTG	GGAGGACATC	600
	CTGAGCGATG	AGGTCAACGT	GGCCCGTGGT	GTGGCCAGCC	TCTTTGCCGG	ACGCTCTGTG	660
35	GCTCTGTGG	ACAGAAAGAG	CTCTGGACCG	TACTCTATCT	ACATTGCCAA	TTAGCGCTAC	720
	GGTAATGTGG	GGCTGTATGC	CCTCATTGAA	ATGGACCTTG	AGGCGAGTGA	CCTCTCCCGG	780
	GGCATTTCTG	GCTCAGAGAG	TGTGGCTGCT	GAGGCTGGGG	TCAGCAATA	TACAGAAGGC	840
	TTCTCCCAACA	CTGCTCTCTC	AAGCATTTGGT	GAGATATCTG	GCAGAACCGA	GGAGCGGAA	900
	GGAGGAGACC	CAGAGGAGGC	AGATGAGGAG	CACAGTGGGG	ATGGAAGCAC	CAGCCAACTG	960
40	TGCTCGCTGG	GCTGGAAGGA	CGGCGAGTTC	AAGGAAGAG	CAGCAGCTTT	GGTGGAGGAA	1020
	CAGAGGAGAG	CTGGGGGAGC	TGGCGTGGCC	AGAGGAGCTG	TTGGAACAGC	TTCTGCAACT	1080
	TCGAAGAGCT	ATTGGCTTGA	CAAGAACCTA	TTTGGCCCA	CATGTTACTA	TTCTGTCTGC	1140
	GGGCTTTCTC	CAGCCACACC	TTTCCCTGCC	CGCCAGGCC	CCCAACACTA	CCCTGTAGCC	1200
	CCCTTTGTCA	CTCAGCTAAT	GACACATGGA	CGTCTGGCTG	GAAGACTAGC	CGGAGGTGTC	1260
45	CCCCACCCCC	GAGCCCCAGG	AATGGACCCC	AAATGTAAGG	CGCCCATGTC	TGAGCCCGGC	1320
	CTGATGGCTG	AGGCTTTGGG	CGCGTGGCCA	GGCTCAGCA	CCACTGTGGT	GGCAGGGGCG	1380
	CTGAGAAAGT	GGGAGGAAAG	CAGGCAGAG	GGGCAGGCCA	TGTCCAGATG	TGCACTCAGG	1440
	GAGCTGGGAG	GTCCCTGGAG	CCAAGCCACA	CAGCACCTGC	CTGCTAGAGA	GCTGTATGAC	1500
	CTGGGAGAAC	CTCCCATTTT	ACAAAGAAC	GACGAGATC	CAGGGAGGAG	AAGGGACTCG	1560
50	CCCAAGGTCA	CACAGGAGTG	CCATCTAGTG	GCCACCATGC	CAGCTCTCGG	GGGACTCGAG	1620
	GGCCCGGGGA	GGGTGGCCAA	GGGAGAGATT	GGGAGAGAGA	CTGGGCGAGT	AGGAAGACCA	1680
	CTCTCCCATC	CCCTGTCTCC	CAACTTCCCC	AGCTGCTTGA	GGCTCTTGA	AGCCGGAGCA	1740
	GTGCGGGGAG	CTGCCCTGCC	TGGGAATCCT	GGGAAGTGGG	TTCTGGACAT	GGCCAGGGCC	1800
	CTGGCGTGGG	ACCAAGATGGA	AAAGAGGAGG	GGGAAGATTC	ATGGAGACCA	TGAGCCACGA	1860
55	TTTAGGCTCA	GGAAAGCAGG	GGAAAGCAGG	TTCCCTCCAG	GCTCTCTTGA	GGAGCCTCTG	1920
	CTGCACTTCC	CCCTAGGCTT	CAGAGGCAGC	CCTGTCTCTC	AGGTGGGCTT	GGGGCTTGCT	1980
	TCTGCCACTC	ACTGTGGGTC	GATGTCTTTT	CTAGGGGGCC	GAGGCGTCHG	CGTGGGCCCC	2040
	ATCCTCAGCA	GCACTGCCCT	GGATATCTTC	TGGGACAAAT	AGAATGGGCT	TAATCTCTCT	2100
	TTCCACAACC	GGGGCGATGG	CACCTTTGTT	GACGCTGCGG	CCAGTGTCTGA	AGTGTGTTTA	2160
60	GGCTTCATCG	TTCACTCAAA	ATATCACTTC	TGCAGAGATT	TTCTCTCACTC	CCTGTGCCAC	2220
	CTAGCAGAAA	CTGGTCTCTC	CTCTCTCTTC	TGCCCGTGGC	ATGCACGTCT	TCTTCAGGCT	2280
	CCACATTTGC	ATCATGGTTT	GTCTATGAGC	TTTACAAGGA	CGGGTTCAGG	GTCTATTTCA	2340
	TTCTTGACGC	AAGGCTTGGC	CTCCAGTGCC	CACCGAGGGA	CACTCAGGCT	CCAGGCTTCT	2400
	CAGGGGGCCC	CACTCTGCTT	TCTGCAAGA	GCTCCCTGTC	TCTGGGGTTC	TCTGATCCCC	2460
65	ACTGCTTATT	ACATTGTCTT	GTGGTCTGCC	ATCCACAGAG	GGCTGATGAC	CCACAGCTAT	2520
	TTGTCTCTTG	AAAGAGTCAA	CGTGGGTGTT	GACGACCCCC	ACCAGCATGG	GGAGGTTGTC	2580
	GGCTGGCTG	ACTTCAACCG	TGATGGCAAA	GTGGACATCG	TCTATGGCAA	CTGGAAATGGC	2640
	CCCCACCGCC	TCTATCTGCA	AATGAGCAC	CTTGGGAAGG	TCCGCTTCCG	GGACATGCC	2700
	TCACCCAGST	TCTCCATGCC	CTCCCTGTTC	CGCAGGTTCA	TCACGCGGGA	CTTTGACAA	2760
70	GACAGGAGC	TGAGATCTT	CTTCAACAC	ATTGCTTACC	GCAGCTCTTC	AGCCCAACCG	2820
	CTCTTCCGAT	GCTCCATCTT	GGCTCTGTC	TCTTCTATCT	TGACAGCTGG	TGGGAGGAAC	2880
	GGTCAGGAG	AAGGTTTAA	AATCAGAAAG	GGAGGTTTCC	CAGGGCCAGG	GGGTCAAGCG	2940
	AAGGTCAACA	CAGGTCTCTT	BATGAAGAAA	CAGAAAGGAA	GGAAAGGAGA	GGACTGCGCA	3000
	AGAGGCTGTG	GGAAATGAGG	GCAAGGCTTG	GCCAGGAGTC	CGGCTCTGTC	TATTGACAGG	3060
75	AAAGGGAGAG	GAATGTGGC	CCAAAGTGTG	CCAGAAACCC	AAGCGCCACA	AGATACAAAG	3120
	CCACACTACC	ACAAAAGGGG	GCTACAGGGT	CCAATCACTA	CCAGGAAAGG	GGGCTACGGG	3180
	GTCCAAATAC	TACCAAGAAA	AGGGGCTHCG	GGGTCCAATC	ACTACCAGGA	AAAGGGGCTA	3240
	CGGGTCCAA	TACTACACAG	GAAAAGGGGC	TACGGGTGCC	AATCACTACC	AGGAAAGGGG	3300
	GCTACGGGCT	CCATCACTA	CCAGGAAAGG	GGGCTACAGG	GTCCAACTAC	TACCAAGGAA	3360
80	AGGGGCTACG	GGCTTCAATC	ACTACCAGGA	AAAGGGGCTA	CAGGGTCCAA	TCACTACCAC	3420
	AGAAAGGGGC	TACGGGCTCC	AATCACTACC	AGGAAAGGGG	GCTACGGGCT	CCAATCACTA	3480
	CCAGGAAAGG	GGGCTACAGG	GTCCAACTAC	TACCAGGAAA	AGGGGCTACG	GGGTCCAATC	3540
	ACTACCAGGA	AAAGGGGCTA	CGGGCTCCAA	TCACTACCAG	GAAAAGGGGC	TACGGGGTCC	3600
	AATCACTACC	AGGAAAGGGG	GCTACAGGGT	CCAATCACTA	CCAGGAAAGG	GGGCTACAGG	3660

GTCCAATCAC TACCACAGAA AGGGGCTACG GGCTCCAATC ACTACCAGGA AAAGGGGCTA 3720
 CGGGGTCCAA TCACCTACCAG GAAAAGGGGC TACGGGCTCC AATCACTACC AGGAAAAGAG 3780
 GCTATGGGGT CCAATCACTA CCAGGAAAAG GGGCTACGGG CTCCAATCAC TACCAGGAAA 3840
 AGGGGCTATG GGGTCCAATC ACTACCACAG AAAGGGGCTA CGGGGTCCAA CGTCATCCGT 3900
 AGAGAGCAGC GAGACCCCTT CATCGAGGAG CTCATATCCG CGGACGCCCT GGAGCCTGAG 3960
 GCGCGGGGCA CAGGGGGTGT GGTGACCGAC TTCGACGGAG ACGGGATGCT GGACCTCATC 4020
 TTGTCCCATG GAGAGTCCAT GGCTCAGCCG CTGTCCGTCT TCGGGGGCAA TCAGGGCTTC 4080
 AACCACAACT GGCTGCGAGT GGTGCCACGC ACCCGGTTTG GGGCTTTTGC CAGGGGAGCT 4140
 AAGGTGCTGC TCTACACCAA GAAGAGTGGG GCCCAGCTGA GGATCATCGA CGGGGGCTCA 4200
 GGCTACTCTG TGGAGATGGA GCCCGTGGCA CACTTTGGCC TGGGGAAGGA TGAAGCCAGC 4260
 AGTGTGGAGG TGAAGTGGCC AGATGGCAAG ATGGTGAGCC GGAACGTGGC CAGCGGGGAG 4320
 ATGAATCTAG TACTGGAGAT CCTCTACCCC CGGGATGAGG ACACACTTCA GGACCCAGCC 4380
 CCACTGGAGT GTGGCCAAGG ATTCTCCAG CAGGAAAATG GCCATTGCAT GGACACCAAT 4440
 GAATGATCC AGTTCCTATT CGTGTCCCTT CGAGACAAGC CGGTATGTGT CAACACCTAT 4500
 GGAAGCTAGA GGTGCGGAGC CAACAAGAAG TGCAGTCCGG GCTACGAGCC CAACGAGGAT 4560
 GGCACAGCCT GCGTGGGTAC TGAGCTAGGC TCTAGGCATA CAATGACGTG GAAACCAAGG 4620
 CCAAAAAGAG ACCTGCAACT TTCCAAGGC ATCTGCACCC CGGTCTGGTC CTTTTCCTG 4680
 CCGGTTGGCC GGTCTCTCCT CAAAAGAGCT CAGCTCCAGG CTGCTCCAG CACCTTCTC 4740
 CAGAAAGCTC CAGGTATTCC ARAAGCCCAA GTGTATGAAC AAGATCAGGA ATAA

Seq ID NO: 185 Protein sequence

Protein Accession #: FGENESHH

1 11 21 31 41 51
 MACPGGLEAR CAGWMLGPGP GSSSPASPPH SSSRYNGPNL VLKYDRAQKH LVNIAVDERS 60
 SPYYALDRDQ GMAIGWTACD IDGDGKEELY PLNTNNAPSG HSSSAQVPSG IERNRFVLKP 120
 PPTTPAGLLG LPPLSGRDPF SSIQASPDFS RQGERVPVPC CROGLRPTHE PEPPFLRPKS 180
 GVATYTDKLF KFRNNRWEDL LSDEVNVARG VASLEAGRSV ACVDRKSGSR YSIYIANYAY 240
 GNVGFDALIE MDEPESDLER GILALRDVAA EAGVSKYTEG FSHITASPSIG EISGRTEERE 300
 GGFPEADEEE HSGDGSSTQL CRIGWKDGQF KEEBAALVEE QREAGAAGVP RGRVRTALQT 360
 SKSLADKNL FGPFCYISVC APSPAHFFFA RQAPQHYEVA PLVTQLMTHG RLAGKLARSV 420
 PHPRAPGMDP KCKGRHASPQ LMAEALGAWP ALSTTVVPGG LRSWERSRQK QJAMERCALR 480
 ELGGPWSQAT QHLPARELYD LGEPPILQRT DGDGRRRDS PKVTQECHEV ATMPALGGLE 540
 GPGRVAKREI GRETGAVGRP LSEPLVNPFP SCLRFLEAGT VPGAALFQNP GNVLDMAKA 600
 LAMNQMEKEE CKIHGHEHPR FRLRKARBAE FPPGSSSEPL LQFPSSGLRGS PVLQVGLGLA 660
 SATRCGSMSE LGGRGVSVGP ILSSASDIF CDNENGPNEL FHNRGDGTFFV DAAASAEREL 720
 AFIVHLKYHL CRDFFPHSLCH LAETGPSSSC CFWHARLQA PHCHHGLEMS FTRTGSRFYS 780
 FLTQGLASSA HRRTLSLQGS QGAPPLCLAR APCVLGSLIP TAYYIWIWSA IPBSLMTHSY 840
 LSSERVNVGV DDPHQHGRGV ALADFARDGK VDIVYGMWNG PHRLYLQMSI HGKVRFDIA 900
 SPKFSMPSPV RTVITADFDN DQELSLFFPN JAYRSSSANR LFRCSILARS SSSLTAGGEN 960
 GQGBELRIRR GGFPGSGQA KMTGSLMKK QKGRKEDWA RGCNAGQSL AKEPASAIAG 1020
 KGKGNVAQSV PRTPAQPDTK PHYHKKGLQG PITTTRKQYG VQSLPGKGAT GSEHYQEKEL 1080
 RGFITTRKRG YGVQSLPGKG ATGSNRYQEK GLQGFITTRK RGYGLQSLPG KGATGSNHYH 1140
 RKGLRAPITT RKGQVGVQSL PGKAGTGSNH YQEKGLRQPI TTRKRGYGLQ SLPGRGATGS 1200
 NHYQEKGLQG PITTTRKRGYR VQSLPGKGAT GSNHYQEKEL RGPITTRKRG YGLQSLPGKE 1260
 AMGSNRYQEK GLRAPITTRK RGYGVQSLPO KGATGSNVR BEHGDPLIEE LNFQDALEPE 1320
 GRGTGGVVD FDGDMGLDLI LSHGSEMAQF LSVFRGNQGF NNNWLKVVPF TRFGAFARGA 1380
 KVVLYTKKSG ARLRIDGGS GYLCEMEPVA HFGLGKDEAS SVEVTWPDGK MVSERNVASGE 1440
 MNSVLEILYP RDEDTLQDPA PLRCGQGFSPQ QENGHCMDTN ECIQPPFVCP RDKPVCVNTY 1500
 GSYRCRTNKK CSRGYEENED GTACVGTIELG SRHTMTWKPR PKRELQLSQG ICTFVHSFFL 1560
 PGRILLKRA QLQAPSTLL QKAPGTFEAQ VYEQDQE

Seq ID NO: 186 DNA sequence

Nucleic Acid Accession #: NM_000584.1

Coding sequence: 75..374

1 11 21 31 41 51
 AGCAGAGCAC ACAAGCTTCT AGGACAAGAG CCAGGAAGAA ACCACCGGAA GGAACCATCT 60
 CACTGTGTGT AAACATGACT TCCAAGCTGG CCGTGGCTCT CTGGGCAGCC TTCTCTGATT 120
 CTGCAGCTCT GTGTGAAGGT GCAGTTTTCG CAAGGAGTGC TAAAGAACTT AGATGTCAGT 180
 GCATAAAGAC ATACTCCCAA CCTTTCACCC CCAATTTTAT CAAAGAACTG AGAGTGATTG 240
 AGAGTGGACC ACACCTGGCC AACACAGAAA TTATTGTAAA GCTTTCTGAT GGAAGAGAGC 300
 TCTGTCTGGA CCGCAAGGAA AACTGGGTGC AGAGGGTTGT GGRGAGATT TGAAGAGGG 360
 CTGAGAACTC ATAAAAAAT TCAATCTCTG TGGTATCCAA GAATCAGTGA AGATGCCAGT 420
 GAAACTTCAA GCAAACTTAC TTCAACACTT CATGTATTGT GTGGGTCTGT TGTAGGGTTG 480
 CCAGATGCA TACAAGATTC CTGTTAAAT TTGAATTCA GTAAACAATG AATAGTTTFT 540
 CATGTATCCA TGAAATATCC AGAACATACT TATATGTAAA GTATTATTTA TTGAAATCTA 600
 CAAAAACAA CAATAATTT TTAATATATA GGATTTTCCT AGATATTGCA CGGGAGATA 660
 TACAAATAGC AAAATTGAGC CAGGGGCCAA GAGAATATCC GAACTTAAT TTCAGGAATT 720
 GAATGGGTTT GTGTGAAGT GATATTGAA GCATCACATA AAAATGATGG GACATAAAT 780
 TTTGCCATAA AGTCAAATT AGCTGGAAT CCTGGATTIT TTTCTGTAAA ATCTGGCAAC 840
 CTAAGTCTGC TAGCCAGGAT CCACAAGTCC TTGTTCACCT GTGCCCTTGT TTCTCTTTA 900
 TTCTAAGTG GAAAAGTAT TAGCCACCAT CTACCTCAC AGTGATGTTG TGAGGACATG 960
 TGGAGACACT TTAAGTTTT TCAATATAAC ATAAATATF TTCAAGTGA ACTTATTAA 1020
 CTATTATTA TTATGATTT TATTAAACA TCAATATTT GTGCAAGAA TTGGAATAAT 1080
 AGAAGATGAA TCATTGATTG AATAGTTATA AAGATGTTAT AGTAAATTTA TTTTATTTTA 1140
 GATATTAAAT GATGTTTTAT TAGATAAAT TCAATCAGGG TTTTATGATT AAACAAGAA 1200
 ACAATTGGGT ACCAGTTTAA ATTTTCATT CAGATAAACA ACAATATATT TTTTAGTATA 1260
 AGTACATTAT TGTTTATCTG AAAATTTTAA TTGAACATAA AATCCTAGTT TGATACTCCC 1320
 AGTCTTGTCA TTGTCAGCTG TGTGGTAGT GCTGTGTTGA ATTACGGAAT AATGAGTTAG 1380
 AACTATTAAA ACAGCCAAA CTCCACAGTC AATATTAGTA ATTTCTTGCT GGTGAAACT 1440
 TGTATTATAT GTACAAATAG ATTCCTTATA TATTATTATA ATGACTGCAT TTTTAAATAC 1500

AAGGCTTAT ATTTTAACT TTAAGATGTT TTTATGTGCT CTCCTAAAT TTTTACTGT 1560
 TTCGTATGT ATGGAATAT AAAAGTAAAT ATGAACATT TAAATATAA TTTGTTGTA 1620
 AAGTAAAAA AAAAAAAA

5 Seq ID NO: 187 Protein sequence
 Protein Accession #: NP_000575.1

1 11 21 31 41 51
 10 MTSKLAVALL AAPLISAALC EGAVLPRSAK ELRCQCIKTY SKPFHPKFIK ELRVIESGPH 60
 CANTEIIVKL SDGRELCCLDP KENWVQRVVE KFLKRAENS

15 Seq ID NO: 188 DNA sequence
 Nucleic Acid Accession #: NM_003661.1
 Coding sequence: 1..1152

1 11 21 31 41 51
 20 ATGAGTGCAC TTTTCCTTGG TGTGGGAGTG AGGGCAGAGG AAGCTGGAGC GAGGGTGCAA 60
 CAAACCGTTC CAAGTGGGAC AGATACTGGA GATCCTCAAA GTAAGCCCTT CGGTGACTGG 120
 GCTGCTGGCA CCAATGGACC AGAGAGCAGT ATCCTTATTG AGGATGCCAT TAAGTATTTT 180
 AAGGAAAAAG TGAGCACACA GAATCTGCTA CTCCTGCTGA CTGATAATGA GGCCTTGAAC 240
 GGATTCGTGG CTGCTGCTGA ACTGCCCAGG AATGAGGCAG ATGAGCTCCG TAAAGCTCTG 300
 25 GACAACTTGG CAGACACAAAT GATCATGAAA GACAAAAACT GGCACGATAA AGCCACAGCAG 360
 TACAGAAACT GGTTCCTGAA AGAGTTTCCT CGGTTGAAAA GTGAGCTTGA GGATAACATA 420
 AGAAGGCTCC GTGCCCTTGC AGATGGGGTT CAGAAGGTCC ACAAGGCCAC CACCATCGCC 480
 AATGTGGTGT CTGGCTCTCT CAGCAATTTC TCTGGCATCC TGACCTTCGT CGGCATGGGT 540
 CTGGCACCTT TCACAGAGGG AGGCAGCCTT GTACTCTTGG AACCTGGGAT GGAGTTGGGA 600
 ATCAGACCGG CTTTGACCGG GATTACACAG AGTACCATGG ACTACGGAAA GAAGTGGTGG 660
 30 ACACAGCGCC AAGCCCAACG CTTGGTCATC AAAAGCCTTG ACRAATTGAA GGAGGTGAGG 720
 GAGTTTTTGG GTGAGAACAT ATCCAACCTT CTTTCCTTAG CTGGCAATAC TTACCAACTC 780
 ACACGAGGCA TTGGGAAGGA CATCCGTGCC CTCAGACGAG CCAGAGCCAA TCTTCAGTCA 840
 GTACCGCATG CCTCAGCCTC ACGCCCCCGG GTCACGTGAGC CAATCTCAGC TGAAGCGGCT 900
 GAACAGGTGG AGAGGGTTAA TGAACCCAGC ATCCTGGAAA TGAGCAGAGG AGTCAAGCTC 960
 35 ACGGATGTGG CCCCTGTAAAG CTTCTTTCTT GTGCTGGATG TAGTCTACCT CGGTACGAA 1020
 TCNAAGCACT TACATGAGGG GGCAGAGTCA GAGACAGCTG AGGAGCTGAA GAAGGTGGCT 1080
 CAGGAGCTGG AGGAGAAAGCT AAACATCTCT AACAATAATT ATAAGATTCT GCAGGCGGAC 1140
 CAGAACTGT GA

40 Seq ID NO: 189 Protein sequence
 Protein Accession #: NP_003652.1

1 11 21 31 41 51
 45 MSALFLGVGV RAREAGARVQ QNVPSGTDTG DPQSKFLGDW AAGTMDPSS IFIEDAIKYF 60
 KEKVSTQNL LLLTDNEAWN GFVAAAELEP NEADELRKAL DMLARQMIMK DKNWHDKGOQ 120
 YRWVFLKEFF RLKSELEDNI RRLRALADGV QKVEKQTTIA NVVSGSLGIS SGILTLVGMG 180
 LAFPTGGSL VLLBEGMBEL ITAALTGITS STMDYGGKHW TQAQANDLVI KSLDKLKEVR 240
 50 EFLGEMISNF LSLAGNTYQL TRGIGKDIRA LRRARANLQS VPHASASRPR VTPEISAESG 300
 EQVERVNEPS ILEMERGKVL TDVAPVSFFL VLDVVYLVEY SKHLREGRAS STAEELKKVA 360
 QLEERKLNL NNNYKILQAD QEL

55 Seq ID NO: 190 DNA sequence
 Nucleic Acid Accession #: NM_014452.1
 Coding sequence: 1..1968

1 11 21 31 41 51
 60 ATGGGGACCT CTCGAGCAG CAGCACCGCC CTCGCTCCTT GCAGCCGCAT CGCCCGCGA 60
 GGCACAGCCA CGATGATGCG GGGCTCCCTT CTCCTGCTTG GATTCTCTAG CACCACACA 120
 GCCTCAGCCAG AACAGAAGGC CTCGATCTCT ATTTGGCAGT ACCCGCATGT TGACCGTGCC 180
 ACCCGCCAGG TGCTAACCTG TGACAGAGGT CACGACAGAA CCTATGCTCT TGAGCATTGT 240
 ACCAACACAA GCGTGGCGGT CTGACGAGT TGCCCTGTGG GAGACCTTAC CAGGCATGAG 300
 65 AATGGCATAG AGAAATGCCA TGACTGTAGT CAGCCATGCC CATGGCCAAT GATTGAGAAA 360
 TTACCTTGTG TCGCTTGAC TGACCGAGAA TGCACTTGCC CACTGGCAT GTTCCAGTCT 420
 AAGCTACCT GTGCCCCCA TACGGTGTGT CCTGTGGGT GGGGTGTGCG GAAGAAAGGG 480
 ACAGAGACTG AGGATGTGCG GTGTAGCAG TGTGCTCGGG GTACCTTCTC AGATGTGCTC 540
 TCTAGTGTGA TGAATGCAA AGCATAACA GACTGTCTGA GTCAGAACCT GGTGGTGATC 600
 AAGCCGGGGA CCAAGGAGAC AGACAACGTC TGTGGCACAC TCCCTCTCTT CTCAGCTCC 660
 70 ACCTCACCTT CCCCTGGCAC AGCCATCTTT CCAAGCCCTG AGCAGATGGA AACCCATGAA 720
 GTCCTTCTCT CCACTTATGT TCCCAAGGC ATGAACCTCA CAGAAATCCA CTCCTCTGCC 780
 TCTGTAGAG CAAAGTACT GAGTAGCATC CAGGAAGGGA CAGTCCCTGA CACACACAGC 840
 TCAGCAAGGG GGAAGGAGA CGTGAACAG ACCCTCCCAA ACCTTCAGGT AGTCAACCAC 900
 CAGCAAGGCC CCCACACAG ACACATCTCT AAGCTGTGCG CGTCCATGGA GGCCTCTGG 960
 75 GGCGAGAGT CCAGCAGGCG CATCAAGGCG CCAAGAGGG GACATCTAG ACAGAACCTA 1020
 CACAGCAAT TTGACATCAA TGAGCAATTG CCCTGGATGA TTGTGCTTTT CCTGCTGCTG 1080
 GTGCTGTGG TGATTGTGGT GTGAGTATC CGGAAAGCT CGAGGATCTT GAAAAGGGG 1140
 CCCCGGCAGG ATCCAGTGC CATTTGTGAA AAGGCAGGGC TGAAGAAATC CATGACTCCA 1200
 ACCCAGAAC GGGAGAAATG GATCTACTAC TGCAATGGCC ATGGTATCGA TATCTTGAG 1260
 80 CTTGTAGAC CCCAAGTGGG AAGCCAGTGG AAGATATCT ATCAGTTCTT TTGCAATGCG 1320
 AGTGAGAGG AGGTTGCTGC TTCTCCAAT GGGTACACAG CCGACCAAGA GCGGGCTTAC 1380
 GCAGCTCTGC AGCACTGGAC CATCCGGGCG CCGAGGCCA GCTCGGCCA GCTAATTAGC 1440
 GCCCTGCGCC AGCACCAGG AAGCATGTT GTGAGAGA TCCGTGGCT GATGGAGAC 1500
 ACCACCCAGC TGGAACTGA CAACTAGCT CTCCGATGA GCCCCAGCC GCTTAGCCG 1560

5
AGCCCCATCC CCAGCCCCAA CGCGAACTT GAGAAATCCG CTCTCTTGAC GGTCGAGCCT 1620
TCCCCACAGG ACAAGAACA GGGCTTCTTC GTGGATGAGT CGGAGCCCTC TCTCCGCTGT 1680
GACTCTACAT CCAGCGGCTC CTCGCGCTG AGCAGGAACG GTTCTTTAT TACCAAGAA 1740
AAGAAGGACA CAGTGTTCG GCAGGTACGC CTGGACCCCT GTGACTTGCA GCCTATCTTT 1800
GATGACATGC TCCACTTCT AAATCTGAG GAGCTGCGGG TGATGAAGA GATCCCCAG 1860
GCTGAGGACA AACTAGACCG GCTATTGAA ATTATTGAG TCAAGAGCCA GGAAGCCAGC 1920
CAGACCTCC TGGACTCTGT TTATAGCCAT CTCTCTGACC TGCTGTAG

Seq ID NO: 191 Protein sequence
Protein Accession #: NP_055267.1

10
1 11 21 31 41 51
| | | | |
15 MGTSPSSSTA LASCSRIARR ATATMIAGSL LLLGFLSTTT AQPEQKASNL IGTYYRHVDRA 60
TGQVLCTDKK PAGTYVSEHC TMTSLRVCSS CPVGTFTRHE NGTEKCHDCS QPCFWPMIEK 120
LPCAALTDRE CTCPEGMFQS NATCAPHTVC FVGWGVRRKG TETEDVRCKQ CARSTFSDVP 180
SVVMCKKATY DCLSQNLVVI KPGTKETDNV CGTLPSSSSS TSPSPGTALF PRPEHMETHE 240
VPSTYVPRK MNSTESNSBA SVRPKVLSSI QEGTVEDNTS SARGKEDVNK TLPMLQVVNH 300
20 QQCPHHRHIL KLLPSMEATG GEKSSSTPIK PKRGHPRQNL HGHFDINEHL PWMIVLFLLL 360
VLNVIVVCSI RKSSRTLKKG PRQDPSALVE KAGLKKSMTP TONREKWIYY CNGHGLDILK 420
LVAAQVGSOW KDIYQFLNCA SEREVAESEN GYTADHERAY AALQHWITRG FEASLAQLIS 480
ALBQHRNDV VEKIRGLMED TTQLETDLA LPMSPSFLSP SPIPSNPKL ENSALLTVEP 540
SPQDKNKQFF VDESPLLEC DSTSSGSSAL SRNGSPITKE KNDTVLRQVR LDPCLQPIF 600
25 DDMLEFLNPE ELRVETETPO AEDKLDRLF IIGVKSQEAS QTLIDSVYSH LPDILL

Seq ID NO: 192 DNA sequence
Nucleic Acid Accession #: XM_044533
Coding sequence: 238..2751

30
1 11 21 31 41 51
| | | | |
35 GCTCTGCCA AGCCGAGGCT GCGGGGCGGG CGCGGGCGGG AGGAATGCGG TGCCCCGCGG 60
AGGGGCTGAG TTGGCCAGGG CCCACTTGAC CTTGTTTCCC AACTCCCGCC CCCAGGTCC 120
GGAGGCGGGG GCCCCGCGGG GACTTGGGGG GGGGACCGCG GGGCGGAGCT GCCCGCGGTG 180
AGTCCGGCCG AGCCCACTGA GCGCGAGCGG CGGGACACCG TGGCTCCTGC TCTCGAATG 240
CTGGCGACCG CGATGGGCTT GAGGAGCTGG CTGCGCGCCC CATGGGCGCG GCTGCGCGCT 300
CBGCTACCGC TGCTGCTGCT CTTGCTGCTG CTGCTCCTGC TGCAGCGCGC GCTCCGACCC 360
TGGGCGCTCA GCGCCCGGAT CAGCCTGCTT CTGGGCTCTG AAGAGCGGGC ATTCTCAGA 420
TTGAGAGCTG AACACATCTC CAATACACAC GCGCTTCTGC TGAGCAGGSA TGGCAGGACC 480
40 CTGTACGTGG GTGCTCGAGA GCGCTCTTTT GCACCTAGTA GCAACCTCAG CTTCTGCGCA 540
GGGCGGAGGT ACCAGAGGCT GCTTTGGGGT GCAGACGAG AGAAGAAACA GCAGTGCAGC 600
TTCAAGGGCA AGGACCCACA GCGCGACTGT CAAACTACA TCAAGATCCT CCGCGCGCTC 660
AGCGGCGATC ACCTGTTTAC CTGTGGCACA GCAGCCTTCA GCGCCATGTG TACCTACATC 720
AACATGAGGA ACTTCACTCT GGCAGGGGAC GAGAAGGGSA ATGTCTCTCT GGAAGATGCG 780
45 AAGGGCGGTT GTCCCTTCGA CCGGAATTTC AAGTCCACTG CCTTGGTGGT TGATGGCGAG 840
CTCTACACTG GAACAGTCAG CAGCTTCCAA GGGGAATGACC CGGCCATCTC GCGGAGCCAA 900
AGCCTTCCGC CCACCAAGAC CGAGAGCTCC CTCAACTGGC TGCAAGACCC AGCTTTTGTG 960
GCCTCAGCCT ACATTCTCTG GAGCCTGGGC AGCTTGCAAG GCGATGATGA CAGATCTAC 1020
50 TTTTCTTCA GCGAGACTGG CCAGGAATTT GAGTTCTTTG AGAACACCAT TGTGTCCCGC 1080
ATTGCGGCGA TCTGCAAGGG CGATGAGGGT GGAGAGCGGG TGCTACAGCA GCGCTGGACC 1140
TCTTCTCTCA AGGCCAGCT GCTGTGCTCA GCGCCGACG ATGGCTTCCC CTTCAAGCTG 1200
CTGCGAGATG TCTTCACTCT GAGCCCGGCG CCCGAGACT GCGGTGACAC CTTTCTCTAT 1260
GGGCTCTTCA CTTCOCAGTG GCACAGGGGA ACTACAGAAG GCTCTGCCGT CTGTGTCTTC 1320
55 ACAATGAAGG ATGTGCGAGG AGTCTTTCAG GCGCTCTACA AGGAGGTGAA CCGTGAGACA 1380
CAGCAGTGGT ACACGCTGAC CCACCCGGTG CCCACACCCC GGCCTGAGAC GTGCATCAC 1440
AACAGTGCCG GGGAAAGGAA GATCAACTCA TCCCTGCGAG TCCACAGACG CGTGTGTAAC 1500
TTCTCAAGG ACACATCTCT GATGGACGGG CAGGTCCGAA GCGCGATGCT CTGTCTGAG 1560
60 CCCCAGGCTC GCTACAGCG GGTGGCTGTA CACCGGCTCC CTGGCTGCA CCACACTTAC 1620
GATGTCTCTT TCTTGGGCAC TGGTGACGGC GCGCTOCACA AGGCAGTGAG CBTGGGCCCC 1680
CGGGTGACCA TCATTGAGGA GCTGCAGATC TTCTCATCGG GACAGCCCGT GCAGATCTG 1740
CTCTGGACA CCACAGGGG GCTGCTGTAT GCGGCTCAC ACTCGGCGGT AGTCCAGGTG 1800
CCATGGGCA ACTGCAGGCT GTACAGGAGC TGTGGGACT GCTCTCTGCG CCGGACCCCC 1860
65 TACTGTGCTT GAGCGGCTC CAGCTGCAAG CACCTCAGCC TCTACAGGCC TCAGCTGGCC 1920
ACCAGGCGGT GATCCAGGA CATCGAGGA GCGAGGCGCA AGGACCTTTG CAGCGGCTCT 1980
TCGGTTGTGT CCGCTCTTT TGTACCAACA GGGGAGAGC CATGTGAGCA AGTCCAGTTC 2040
CAGCCCAACA CAGTGAACAC TTTGGGCTGC CCGCTCTCTT CCAACCTGGC GACCCGATCT 2100
TGGTACGCA ACGGGGCCCC GTCAATGCC TGGGCTCTCT GGCACGTGCT ACCACTGGG 2160
GACCTGCTGC TGTGGGCGC CCAACAGCTG GGGGAGTTCC AGTCTGTGTC ACTAGAGGAG 2220
70 GGCTTCCAGC AGCTGGTAGC CAGCTACTGC CCAGAGGTGG TGGAGGACGG GGTGGCAGAC 2280
CAACAGATG AGGTGGGAG TGTACCCGTC ATTATCAGCA CATCGCGTGT GAGTGACACA 2340
GCTGTGGGCA AGGCCAGCTG GGTGCGAGAC AGGTCTACT GGAAGGAGTT CTTGGTGATG 2400
TGACGCTCT TGTGCTGGC CGTGTGCTC CAGTTTAT TCTTGCTCTA CCGGACCCGG 2460
75 AACAGCATGA AAGTCTTCTT GAGCGAGGG GAATGTGCCA GCGTGACCCC CAGACCTGC 2520
CCGTGGTGC TGGCCCTCTG GACCCGCCCA CTCACCGGCC TAGGGCCGCC TAGCACCCG 2580
CTGATACACC GCGGCTACCA GTCCCTGTCA GACAGCCCCC CGGGGTCCCG AGTCTTCACT 2640
GAGTCAGAGA AGAGGCCACT CAGCATCCAA GACAGCTTGG TGGAGGTATC CCCAGTGTGC 2700
CCCGGCGCCG GGTCCGCTT TGGCTCGAG ATCCGTGACT CTGTGGTGTG AGAGCTGACT 2760
TCCAGAGGAC GCTGCCCTGG CTTAGGGGCT TGTGAATGCT CCGAGAGGGT CAACTGGACC 2820
80 TOCCCTCCGC TCTGCTCTTC GTGGAACAG ACCGGTGTGC CCGGCCCCTG GGAGCCTTGG 2880
GGCAGCTGG GCTGCTCTTC TCAGTCAAG TAGCGAAGCT CTTACCCACC AGACACCCAA 2940
ACAGCCGTGC CCCAGAGGT CTTGGCCAAA TATGGGGGCC TGCTAGGTTT GGTGGAACAG 3000
TGCTCTCTAT GTAACTAGAG CCCTTGTGTT AAAAAACAT TCCAAATGTG AAACAGAAAT 3060
GAGAGGGGAG AGATAGCATG GCATGCAGCA CACACGGCTG CTCAGTTTCA TGGCTTCCCA 3120
GGGGTGTCTG GATGCAATCC AAAGTGGTTG TCTGAGACAG AGTTGGAAAC CTTACCAAC 3180

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

TGGCTCTTC ACCTTCCACA TTATCCCGCT GCCACCGGCT GCCCTGTCTC ACTGCAGATT 3240
CAGGACCAGC TTGGGCTGCG TGCGTTCTGC CTGSCAGTC AGCCGAGGAT GTAGTTGTTG 3300
CTGCGCTGCT CCCACCACT CAGGGACCAAG AGGGCTAGGT TGGCACTGCG GCCCTCACCA 3360
GTCCTGCGGC TCGGACCCAA CTCTGAGACC TTCCAGCCT GTATCAGGCT GTGGCCACAC 3420
GAGAGGACAG CGCGAGCTCA GGAGAGATT CTGACAAATG TACGCTTCTC CCTCAGATT 3480
CAGGGAAGAG ACTGTGCGCT GCCTTCCTCC GTTGTGCGT GAGAACCCTG GTGCCCTTC 3540
CCACCATATC CACCTCGCT CCATCTTGA ACTCAAACAC GAGGAACCTAA CTGCACCTG 3600
GTCTCTCTCC CAGTCCCGAG TTCACCTCC ATCCCTCACC TTCTCCACT CTAAGGGATA 3660
TCAACACTGC CCAGCACAGG GCGCTGAAAT TTATGTGTTT TTTATACATT TTTAATAAAG 3720
ATGCACCTTA TGTCTTTTT TAATAAAGTC TGAAGAATTA CTGTTT

Seq ID NO: 193 Protein sequence
Protein Accession #: XP_044533.3

1 11 21 31 41 51
| | | | |
MLRTAMGLRS WLAAPNGALP PRPPLLLLL LLLLLQPPPP TWALSPRISL PLGSEERFPL 60
RPEAEHISNY TALLLSRDRG TLVVGAREAL FALSSNLSPF PGGEYQELLW GADAERKQCC 120
SFKGKDPQRD CONYIKILLP LSGSHLPFCG TAAPSPMCTY INMENFTLAR DEKGNVLEED 180
GKGRCPFDPN FKSTALVVDG ELYTGTVBSF QGNPDAISRS QSLRPTKTES SLNWLQDPAF 240
VASAYIPESL GSLQDDDKI YFFPSETGQR FEFENTIVS RIARICKGDE GGERVLQQRW 300
TSFLKAQLLC SRPDGPPFN VLQDVFTLSP SPQDWRDTLF YGVFTSQHER GTTGSAAVCV 360
PTMKDVQRFV SGLYKEVMRE TQQWYTVTHP VPTPRPGACI TNSARERKIN SSLQLPDRVL 420
NFLKDHFLMD GQVRRMILL LQPARYQRYA VHRVPLHHT YDVLFLGTGD GLHKAVSVG 480
PRVHIIEBLQ IFSSGQPVQN LLLDTHRGLL YAAHSGVVQ VPMANCSLYR SCGDCLLARD 540
PYCANGSSC KHVSLYQPOL ATREWIQDIE GABAKDLCSA SSVVSPSFVP TGEKPCBQVQ 600
PQPTVNTLA CPLLNLRLT LMLRNGAEVN ASASCHVLPF GDLLLVGTQQ LGHFQCHSLE 660
EGFQQLVASY CFEVVEDGVA DQDEGGSVP VVISTSRVSA PAGGKASWGA DRSYWKFLV 720
MCLFLVLAFL LEVFLLLYRH RNSMKVFLKQ GECASVHPKT CPVULFPETR PLNGLGPPST 780
PLDHRGYQSL SDSPPGSRVF TESEKRLPSI QDSFVSVSV CPRPRVRLGS EIRDSVV

Seq ID NO: 194 DNA sequence
Nucleic Acid Accession #: NM_022819.1
Coding sequence: 1..635

1 11 21 31 41 51
| | | | |
ATGGCAGATG GGGCAAAAGC CAACCCCAA GGGTTCAAAA AGAAGGTGCT GGATAGATGC 60
TTCTCTGGGT GGAGGGGCC ACCTTCCGCG GCTCTCTGTC CTCAAGAAC CTCAGGTCT 120
AGCTTGGGTA TGAAGAAGTT CTTCACCGTG GCCATCCTTG CTGSCAGCGT TCTGTCCACA 180
GCTCAGGGCA GCTCTGCTCA CCTGAAGGCC ATGGTGGAGG CGGTACAGG GAGGAGCGCC 240
ATCTGTCTCT TCGTGGGCTA CGGTGCTAC TGTGGGCTGG GGGGCGGTGG CCAGCCCAAG 300
GATGAGGTGG ACTGTGCTCG CCACGCCAC GACTGCTGCT ACCAGGAAGT CTTTGACCAA 360
GGCTGTACCC CCTATGTGGA CCCTATGAT CACACCATCG AGAACACAC TGAGATAGTC 420
TGCAGTGACC TCACACAGAC AGAGTGTGAC AAGCAGACAT GCATGTGTGA CAAGAACATG 480
GTCTGTGTGC TCATGAACCA GAGTACCGA GAGGAGTACC GTGGCTCTCT CAATGTCTAC 540
TGCCAGGGCC CCAAGCCCAA CTGCAGCATC TATGAACCGC CCGCTGAGGA GGTCACTGTC 600
AGTCACCAAT CCCCAGGCC CCGCGCCCT CCGTAG

Seq ID NO: 195 Protein sequence
Protein Accession #: NP_073730

1 11 21 31 41 51
| | | | |
MADGAKANPK GPICKVLDRG FSGWRGPRFG ASCPSRTSR SLMGKKEFTV AILAGSVLST 60
AEBSLLNLKA MVEAVTGRSA ILSFVGYYCY CGLGGRGQPK DEVDNCCAR DCCYQELFDQ 120
GCHPYVDHYD HTIENNTETV CSDLNKTECD KQTCMCDKNM VLCLMNQYTR EBYRGFLAVY 180
CQGTPTNCST YEPPEEETVC SHQSPAPPAP P

Seq ID NO: 196 DNA sequence
Nucleic Acid Accession #: XM_028196.1
Coding sequence: 1315..1791

1 11 21 31 41 51
| | | | |
GGCATTGATG CTGTGTGCGC GTGCGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT 60
GTGTGTGTGT GTCTGAGATC ATGGCAGGCT CCTTTCTGT CTGTCTCTCT GCTCTGCCCC 120
AGACTGCGGG GCTGCGAGAG TGAGGATATC TGGCTCAAC AGCTGCTTAT TCCCGATGGG 180
ATGGCCTGGG CTGGGCCCTT GAGGCCAGGC TGACTTGGAC ATGGCAAGAG GGGTCCAGG 240
CTCTGTGGG CAAGCAGAGG GAGGCCCAA TGTGAGGAA CAGAGTCTCC TGGTGGCTG 300
CTGCTCTCTG GAGCGGGTGG AGTCAGGGAA GAGCTGAGCT GGGGAGTAC CCTGGGCTG 360
GGGTACCCYT AGGCCCATG TAGCACCTGT GTTCCCTGTC CTGTAGGTGA CAGGAGCCAG 420
CCAGCAGCAG TGTCTCTCTT CCCCAGGCC TAGGCAGGCG GGTACAGGGG CCAGCAGCTG 480
CGCCCGCCCC ACCTTCTCTT CCACCCACAT GCGGAGGGT GGCCAGGCAG GCAGGTGGAC 540
GAGTCCAGGC AGCGGCTGAG TCAGTGTGTG TGAATGTTC TGGCGCTCC CAGCTGCACC 600
CTGCCCTTAC CTGCCACCAC CTCACCTTCA TCTCAGGCG CTGCGGCCCT GAGCCCCGTC 660
CAGGAATGCA CCTTTAGCCC AGGCTGCTC AGTGAGCTCC GCGGACAGCC AGCCTGCTC 720
CTCCGCCCAT GACCTGCGAG ACCCTCTGCG GCTTCCAGT TCCGTGGGGC TGCACTGAAC 780
ATGCTCCACC CTCTTGGCTG GCAAAACATG GTGGGCCCA GCTGTGCTG GTGCTGGGGT 840
AGAGGCAAGG AAGTGTATGG ACCGAGAGA TGACACCCC AGGGATGAGA TGGGACCCC 900
AGGCAGGGCC CAGGTCACAG GGCCAGGAG AGAGAGCAG GGAGGAGAG AGCTTCTCTG 960
TGGAGGAGCG ATCTTACAGT GGGGCAAGG GTGCTCTGAG TCCGGGTGAA GGCAGGGAG 1020
AGGCTGCCCA GGCCTGCTT GCTTGGCTGG GGTGGGGGCG TGCTGGGAGG TGGCTGGGAG 1080
GCTGGGCTGT GGCAGCTAAG CTGGAGCTTT GGCAGGGTTC CAGAGGCTCC CTCCTTCAAG 1140

5
 10
 15
 20
 25
 30
 35
 40

```

CCTCCTGCTG CACAGAACCC TGGCCCTTGG CCACCCCGTG CTGCTCCCTT GCCCTGGCAG 1200
ACCCAGCAGT GGCTGCTGCT AGTCAGATGG GGTAGCGGGC AGGGGCGCGA GGGGCCACCC 1260
TCCCAGCTGA CCCAGCCTCC TGGGCGGCTT CTTCACAAAC AGCAGGGTAG AAAGATGGGG 1320
CACCCACCCG TCTCTCCAG TGGCCCGGCC CCAGCTGGCA CCACAGCTAT ACCTGGGCTT 1380
ATTCCAGACC TTGTGCGCCG GACCCCTGTG GAGTTGTGGG ATTCCCAAGA GGGGTGTGGG 1440
GATAACCCAG CCAAGTGGGG GCTGCAGCTG TCCACAGATG CACTCAGCCT GGCTCTACC 1500
CCAGGGCCCC GCTGGGCTCT CATTCGCGCG GCCCTTGGCG CGGGGCTCCT CCTGCTCTCC 1560
TGCTCTCTCT GTGCTGCTTG CTGCTGCTGC CGCCGCCACA GGAAGAAGCC CAGGGACAAG 1620
GAGTCCGTGG GTCTGGGCAG TGCCCGCGGC ACCACCACCA CCCACCTGGT GAGGAGCGGC 1680
TCCTTGCTCA CTCAGTCCAG AGAGGGCTTG AAATCCAGGC TCCAGAGCCC AGGGCAGCGA 1740
GGCGAGTTCA GCGCCAGGGA TGGTTTAACC CCCACAGAGG CAGGGCGTTG AGGACCTTCC 1800
TGGCAGGGAA AGTGGGTGAA CAGAGGTGAG AAGGAGGCCA TGCAACAGGG GCTGCCCCAT 1860
GGGCGCGAGG GAGCCACAGC GGGTTCTTGA GGAAGGCAGG GGGTACCCCA GATGCCAGT 1920
TTTGGGTGGG TTGGCCGGGT CTCACAGAGC GAAGCCGAGC ATTTGTGCTT GTTGGGTGGC 1980
CTGGGCTGGA GCGCGGGGGT CTGACCCCAT GTCATGCAAG GGTCTGCCCG GAGCCAGGG 2040
CTCTGATGAG AGCAGTGTTC AGCACCACCT GCCCTTGTCT CCAACTCACT CCAGGTGCAA 2100
CCTGATGTGG ATGGCTCGGA GTCCAGCCCG GGGGATGCTC AGCAATGGGG GTGCCCTGAG 2160
CTCTCCCTGG AGTTGAGATT TGGAAAGCCAG GAGGTGAAGG GCGCCGCTGC GCAGGACCAG 2220
CGGTTCTGCG AGTTTCCGGA AAGGTTGACG GGGGAAGGGC AGACCCCATG CCCTGGGTGG 2280
TGGGAGAGCTG ACAGGGCAGG GGGCCTTGGC TGAGGCCACC CGCTGCGCTC CCAGATCAGG 2340
GTGGGCTTGA GCGCCAGGAG CGACCTGAGG CTTGGGGGCA CGTGGACCC CTATGCGCGG 2400
GTACAGCTCT CCACCCAGGC CGGACACAGA CATGAGACAA AAGTGCACCG AGGCACGCTC 2460
TGCCCGGTGT TTGACGAGAC CTGCTGCTTC CAGGTGAGTC AGGGATGGTC GGCTGGGTGG 2520
GCTTGACAGG CTGGATGGGC CTGGGCTGGG TGGGCTGGG CAGCTGGGTG GGCTGGGCA 2580
GCTGGGTGGG CTTGAGCTAG GGCAGCAGGG CCTGGCTCAC GCGCTGCTCT CAGATCCCGC 2640
AGGCGAGCTT GCGCGGGGCC ACCCTGCAAG TGCACTTTT CAACCTCAAG CGCTTCTCGG 2700
GGCATGAGCC CTTGGGTGAG CTCGCTCTGC CACTGGGCAC CGTGGATCTG CAGCATGTTT 2760
TGGAGCACTG GTACCTGCTG GCGCCGCGCG CTGCCACTCA GGTGAGGTGC TGGTCACAG 2820
GCCACAGCCC AAGCCAGAGC TGGCAGGGAC CTTGCCCTAT GGGCCATCGG AAGACAGGC 2880
CTGATGGGCA GCATTTTCGG GGGTCTGAGC CCCAACTCGG CCAGAATCAC CCTCCGCGGC 2940
TGAAGCCCTT CTGCTGCCCC ACAGCCCGAG CAGGTGCGGG AGCTGTGCTT CTCTCTCGG 3000
TACGTGCCCA GCTCAGGCGG GCTGACCGTG GTGGTGTGGG AGGCTCGAGG CCTGCGTCCA 3060
GGAATTGCGA AGCCCTACGT GAAGGTCCAG CTCATGCTGA ACCAGAGGAA GTGGAAGAG 3120
AGAAAGACAG CACCAAAAAA GGCACGCGG GCCCCTACT TCAATGAGGC CTTCACTTC 3180
CTGGTGGCCT TCAGCCAGGT CCAGAAATGTG GACCTGGTGC TGGCTGTCTG GGCAGCGCAG 3240
CTGCCGCTCC GAACGTAGCC GTAGGCAAG GTGCACCTG GTGCCCGGCG CTGCGGCGAG 3300
CCCTGCGAGC ACTGGGCGA CATGCTGGCC CACGCGCGGC GCGCCATTCG CCAGCGGCAC 3360
CCCTGCGGCG CAGCCAGGGA GGTGGACCG ATGCTGGCCC TGACGCCCCG CCTTCGCTG 3420
CGGCTGCCCT TGCCCCACTC CTGAATBCAC CACATGCCCT TGTCGCCCG CTGAGCCAG 3480
GCCTTGCCCC AGGCGCGCCT GCAGGACCAC TGCAATAAAC GCCTTCTCCT GCC
  
```

Seq ID NO: 197 Protein sequence

Protein Accession #: XP_028196.1

45
 50

```

1 11 21 31 41 51
MGHPFVSPSA PAPAGTTAIP GLIPDIIVAGT ECKLWDSQSG CGENPAKNGL QISTDALSLA 60
STPGPRWALI AGALAAGVLL VSCILCAACC CRRHRKPKPR DKESVGLGSA RGTITTHLVR 120
SGSLLTQSRE GLKSLRLQSPG QRGESFSPEDG LTPTEAGR
  
```

Seq ID NO: 198 DNA sequence

Nucleic Acid Accession #: NM_000612.2

Coding sequence: 553..1095

55
 60
 65
 70
 75
 80

```

1 11 21 31 41 51
TTCTCCCGCA ACCTTCCCTT CGCTCCCTTC CGTCCCCCCC AGCTCCTAGC CTCGCACTCC 60
CTGCCCCCTT CACGCGCGCC CTCCTGCTTT CGCCGAAACC AAGTGGATTA ATTACAGCT 120
TTCTGTTTCT CTCGCTGCTG TTCTCTCCCG CTGTGCGGCT GCGCGCCTCT CGCTGTCTTC 180
TCTGCCCCCT GCGCTCTCTT CGGCCCCCCC CTTTCAAGTT CACTCTGTCT CTCGCATAT 240
CTCTGCCCCC CTCATCTCTT GATACAACAG CTGACCTCAT TTCCGATAC CTTTTCCTCC 300
CCGAAAAGTA CAACATCTGG CCGGCCCCAG CCGAAGACA GCGCGTCTC CCTGSAAT 360
CAGACGAATT CTCCTCCGCC CCGCAAAAAA AAAAGCCATC CCGCGCTCT GCGCCGTGCG 420
ACATTCGCCC CCGCGACTTC GGCAGAGCG GCGCTGGCAG AGGAGTGTCC GGCAGAGGG 480
CCAAAGCCCG CTGTTGGGTT TGCGACACGC AGCAGGGAGG TGGCGCGAG CGTGCAGGC 540
TTCCAGACAC CAATGGGAAT CCGAATGGGG AAGTCGATGC TGGTGTCTCT CACCTCTTTC 600
GCCTTGCCCT CGTGTGCTAT TGCTGCTTAC CGCCCGAGT AGACCTGTTC GCGCGGGAG 660
CTGGTGGACA CCCTCCAGTT GGTCTGTGGG GACCGCGGCT TCTACTTCAG CAGGCCCGCA 720
AGCCGTGTGA GCGCTGCGAG CCGTGGCATC GTTGAAGGAT GCTGTTTCCG CAGCTGTGAC 780
CTGGCCCTCC TGGAGACGTA CTGTGCTACC CCGCCCAAGT CCGAGAGGGA CGTGTGAC 840
CCTCCGACCG TGCTTCCGGA CAACTTCCCC AGATACCCCG TGGGCAAGTT CTTCCAATAT 900
GACACCTGGA AGCAGTCCAC CCAGCGGCTG CGCAGGGGCC TGCTTGCCTT CCTGCGTGCC 960
CGCGGGGGTC AGGTGCTGCG CAAGGAGCTC GAGGCTTTCA GGGAGGCCAA ACGTACCGT 1020
CCCTGATTG CTCTACCCAC CCAAGACCCC GCGCCCCCCC AGAGATGGCC 1080
AGCAATCGGA AGTGAACAAA ACTGCGGCAA GTCTGCAGCC CGCGGCCACC ATCTGTGAGC 1140
CTCTCTCTGA CCAAGGACGT TTCCATCAGG TTCCATCCCG AAAATCTCTC GGTTCACGT 1200
CCCTCGGGG CTTCCTCTGA CCCAGTCCCC GTGCCCGGCC TCCCGGAAAC AGGTACTCT 1260
CTCTCGCCCC CTCCATCGGG CTGAGGAAGC ACAGCAGCAT CTTCAAACAT GTACAAAATC 1320
GATTGGCTTT AAACACCCCT CACATACCTT CCCCCC
  
```

Seq ID NO: 199 Protein sequence

Protein Accession #: NP_000603.1

1 11 21 31 41 51

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

AAAGTTCCCA GCCAATAGAC AGCATGAATC AAGGAACITG CATATATGT GCTCTGAAT 1680
 CTGTTGTCTC CATGGACCAT TCCTCGAGT AGTGGTGAGA TGGCCTGGG TTGCCCTTGG 1740
 CTCTCTCTCC CTCTACTCAG CCTTAAAAAG GGCCTCTTGG AACTTTACCA GCAGCCTCAG 1800
 CTTTACAAAT GCCTTGGTAT GTACCTCTGG CAAATGCCCT GGTATGTACC TCTGGCAAAT 1860
 GCCCCACCTT GGTGATGTG CAACCTTTCC TTCTGCTAGG GTGTACACCT AGCCTGTGCA 1920
 GGTGTACGCC CTGCTAGGGA GTCACTGTAC ACACAACTC TACTGGAAAT CCTGCCAAAC 1980
 TCTGTACCCC TGCAGCTCCT TTACAGTTCA ATCCAATGAT AGAAACCATC CCTTCCCTTT 2040
 CTCCCTTGGC TGTTCACCCA GCCATTCCCT GAAGGCCCTA CCAACAGGAA TATCCAAGAA 2100
 GCTGTGTCTC TCTCTCGAAC CCTGACCAGA TCATCAGCCA CTGAGGCCAG TGGAAATTTCC 2160
 CCAGGCCTTG TTAATAAAAA AAAAAAATA 2190

Seq ID NO: C66 DNA Sequence

Nucleic Acid Accession #: NM_014459.2

Coding sequence: 738..3407

15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80

1 11 21 31 41 51
 GTAGATGCG TCCGCCGCG CCGCTGCCCT AGCCAGCAAT GCAAGATTAG ATCTCTAAAT 60
 GCAGCAAAAC ACTGCTCTGA AACAGACCGG CCGCGCAGCG AAGCAGACAT TTCACGGTGC 120
 GCTGGGGAAG CTTCAAAATA TATCTGTGAC TCTGTCTTCG TTGCTCTTCA TCCCATCAA 180
 TTTCATCAGC GGAGCGGAGC AGCAAGTAAG AATTTCACTT TCGGATCTGC CTAGAGACAC 240
 ACCTCCCTGC TCCCTCCGCC ACTCGATGTG AAGAGTATTC CGGAGTCTCC GGGCGGGAGT 300
 AGATTGTGCG CACCTTAGCG GAGCGAGGGA AAACCTACTG ATTCTTTAGC TCATTATCAT 360
 CTCTCCGAGA CGAGATTTCG TTCTTATCGC CTGCTCATC GCTCAAGTTT GAGCCTCCCG 420
 AAGTCGGGCG GAGAGAGACG AAACCCCTCG CTCACCCCA GCGCGCAGAA GCCACCGCCT 480
 TGCTCCAAGC CCCTGCAGCT CTGCTGCACC GCAGCTCTCT ACCAGTGCAG GATGCTGTAG 540
 ATCAACAGGT TACGAGGACT TGAGCAGAAI AAGGAGAGAC CACCGGGTGC CGCAGCTCGG 600
 GTGAGAGGGG AAAAAAGGAC CCATAGACTT GTGCTCGCG TCGCGCGCGC ACGCTGCGCC 660
 AGGGCCCGAG GCTCGGCGCG ACTCCCTCTC TGGCTCCTCC AGTCCGATG CTCTGCCCC 720
 CACCTTACAG GCTCTGAGT TACCTTTCCA TCTGTGCTG CTTTCTTCTA TGGGCCCTCG 780
 CCCTCACTCT CAAGAACCTC AACTACTCGG TGCGGAGGGA GCAAGGGGCC GGCACGGTGA 840
 TCGGGAACAT CGCAGCGGAT GCTCGACTGC AGCCTGGGCT TCGGCTTGA GAGCGGGG 900
 GCGGAGGGCG CAGCAAGTCT GGTAGCTACC GGTGCTGGA GAACCTCGCA CGCACCTGCG 960
 TGGAGCTGGA CAGCAGACAG GGGCTCTCT ACACCAAGCA GCGCATCGAC CGCGAGTCCC 1020
 TGTGCGCGCA CAATGCCAAG TGCCAGCTGT COCTCGAGGT GTTCGCGAAC GACAAAGAGA 1080
 TCTGCTATAT CAAGGTAGAG ATCCAGGACA TCAACGACAA CGCGCCCTCC TTCTCTCGG 1140
 ACCAGATCGA GGTGAGCTC TCGGAGAAAG CTGCTCCGG CACCGCTTC AACCTACCA 1200
 GCGCAGATGA CCGGACGCC GCGGAGAAAG GGTCTCCGAC CTAAGCTGTC ACGGCGAAG 1260
 ATCAAGGCCT CTCTGGACTG GAGGTTAAGT CCGCGGCGGA CGGCACCAAG TTCACAGAAC 1320
 TGTCTATCCA GAAGGCTCTG GACCGCGAGC AACAGATCA CCATACGCTC GTGCTGACTG 1380
 CCCTGGACCG TGGCGAGCCT CCAAGTTCGG CCACCGTACA GATCAACGTC AAGGTGATTG 1440
 ACTCCAACTA CACAGGCCCG GTCTTCGAGG CGCATCTTA CTTGGTGGAA CTGCCGAGA 1500
 ACGCTCCGCT GGTGACAGTG GTCATGATC TGAACGCCAC CGACGCGAAT GAAGGTCCCA 1560
 ATGGTGAAGT GCTCTACTCT TTCAGCAGCT ACGTCCCTGA CCGCGTGCAG GAGCTCTTCT 1620
 CCATGAGCCC CAGAGACCGC CTAATCCGTC TGAAGGGCAA TCTGGACTAT GAGGAAAAAG 1680
 GGAATGCTGA GATTGACGTG CAGGCGCGAG ACCTGGGGCC TAACCTATTC CCAGCCCACT 1740
 GCRAAGTCA GGTGAGCTC ATCGACGCA ACGACATGC GCGCTCCATC GGTTCGTCT 1800
 CCGTGCGCCA GGGGGCTCTG AGCGAGGCCG CCGCTCCCGG CACCGTCTAT GCGCTGGTGC 1860
 GGGTCACTGA CCGGACTCT GGCAGAACCG GACAGCTGCA GTGTGCGGTC CTAGGCGGAG 1920
 GAGGAGCGGG CCGGCTGCGG GCGCTGCGCG GCGCGGGGG TTCCCTCCCT TTCAAGCTTG 1980
 AGGAGAACTA CGCAACTTTC TACACGCTGG TGAAGTGAAG CCGCTGCGAC CGCGAGACAC 2040
 AAGACGATGA CAGCTGACCT ATCGTGGCGC GGGAGCGGGG CTCTCTCTCC CTCACCTCCA 2100
 CCAAGTCTGT CCGATCAAG ATTCTAGAGC AGAAGCAGAA CCGGCTCTGG TTCACCAAG 2160
 GGCCTCACT GCTTCAGGTG CAGGAGAAC AATCCCGGG AGAGTACCTG GGTCTGTGTC 2220
 TCGCCAGGGA TCCCGAGCTG GCGCAGAGCG GCACCGTATC CTACTCTATC CTGCCCTCGC 2280
 ACATCGGCGA CGTGTCTATC TACACCTATG TGTCTGTGAA TCCACGAAAC GGGGCCATCT 2340
 ACGCCCTGCG CTCTCTTAAC TCGAGCAGA CCAAGGCTTT TGAATTCAG GTGCTGTCTA 2400
 AGGACTCGGG GCGCGCCGCG CACTTGGAGA GCAACGCCAC GGTGAGGGTG ACAGTGTCTAG 2460
 ACGTGAATGA CAACGCCCCA GTGATCCTGC TCCCGAGCTC GCAGAACGAC ACGCGGAGC 2520
 TGCAGGTGCC GCGCAACGCT GGCCTGGGCT ATCTGGTGAG CACTGTGCGC GCGCTAGACA 2580
 GCGACTTCGG CGAGAGCGGG CGTCTCACT ACAGATAGCT GGAAGGCAAC GACGACCAAC 2640
 TGTTTGAGAT CGACCCGCTC AGCGCGGAGA TCCGCACTCT GCACCTTTTC TGGAGGAGCG 2700
 TGACGCCCGT GGTGGAGCTG GTGGTGAAGG TGACCGACCA CGGCAAGCCT ACCCTGTCCG 2760
 CAGTGGCCAA GCTCATCATC CGCTCGGTGA GCGGATCCCT TCCGAGGGGG GTACCAAGGG 2820
 TGAATGGCGA GCAGCACCAC TGGGACATGT CGTGGCGCT CATCGTGAAT CTGAGCACTA 2880
 TCTCCATCAT CTTCTTAGCG GCGATGATCA CCAATCGCGT CAGTGTCAAG CGCGAGACA 2940
 AGGAGATCCG CACTTACAAC TCGCGCATCG CCGAGTACAG CCACCGCGAG CTGGGTGGGG 3000
 GCAGGGGCAA GAAGAGAGAG ATCAACAAAA ATGATATCAT GCTGGTGCAG AGCGAAGTGG 3060
 AGGAGAGGAA CGCATGAAC GTCATGAGC TGGTGAAGC CCGCTCCCTG GCGACCTTCC 3120
 CCATGTACTT CGACTACGAG ACCCGCTTGC CCGTCACTC GCGCGGCTCG GAGGTGATGT 3180
 ATCTCAAACC GCGCTCCAAC AACCTGACTG TCCCTCAGGG GCACGCGGGC TGCCACACCA 3240
 TCTTCAACGG CAGAGCGACT AATGCAAGCG AGAAGCTTGC CACTCGGATC TCCATAATTC 3300
 AGACAGACAA TTTTCCCGCA GAGCCCAATT ACATGGGCG CAGGCAGCAG TTTGTCAAAA 3360
 GTATTTCACT AGCTCCAGCT TTAAGGACCC AGAAAGAGCC AGCCTGAGAG ACAGTGGGCA 3420
 CGGGGACAGT GAGTCAAGCT ACAGTACCA AGACACTAAC AAGGGCTCTCT GCTGTGACAT 3480
 GTCTGTAGG GAGGCACTCA AGATGAAAAC TACTTCAACT AAAAGCCNAC CACTTGAACA 3540
 AGAAACAGAA GAGTGTGTTA ATGTCACAGA TGAATGCGGA GTGCTTGGTC ATTCTGACAG 3600
 GTGCTGGATG CCACAGTTCC CTGCGACCAA TCAGGCTGAA AATGCAGATT ACGCACACAA 3660
 TCTCTTTGTA CCTACAGTTG AAGCTAATGT TGAGACTGAG ACTTACGAAA CTGTGATATC 3720
 CACTGGGAAA AAGACTTTT GTACATTGG AAGAGACAAG CGAGAGCACA CTATTCTCAT 3780
 TGCCAAAGTT AAACTTATTT TAAAGGCCAA ACGTCCCGCT AGCCTCTCTC TCCAGAGGT 3840
 CCGCTCAGCA TCAAGCAGCC CAACCAAGGC GTGCATCGAG CCGTGCACCT CAACAAAAGG 3900
 CTCCCTGGAT GGTGTGAGAG CAAACACAGG AGCCCTGGCT GAAGCAAGCA GTCACTACTT 3960
 GCGCACTGAC AGTCAATATC TGTCACTAG TAAGCAACCA AGAGACCTTC CTTTCTATGGC 4020

TCAGCCACC CTGGGTCG GGCACACTCT CGGAGGGTGC TGCCTCCCTC TGCCCTTCAA 1080
 TCGTTCGCTT GATGGTCCTT GTCAATTCACG CGGGTGCCTG TGTGTGTATG TCTGTGTATG 1140
 TTAAGGGGAA AGAAGGGATC CCAACTGTGT CCTTATCTG TTTTCTACCT CCTCCTTTGT 1200
 TTAATAAAGG CTGAAGCTTT TTGT

Seq ID NO: 203 Protein sequence
 Protein Accession #: NP_004208

1 11 21 31 41 51
 MAQKENSIFY PYGQQTAPSG LSTLPQEVLE KEPVTPSALV LMSRSNVQPT AAPGQKVMEN 60
 SSCTPDILTR HFTIDDFEIG RPLGKKGKFGN VYLAREKKSH FIVALKVLFK SQIEKBSVEH 120
 QLRREIEIQA HLHHPNLLRL YNYFYDRRI YLILEVAPRG ELYKELOKSC TFDQRTATI 180
 MEELADALMY CHGKKVIHRD IKPENLLGL KGLKLIADFG KSVHAPSLRR KTMCGTLDYL 240
 PPEMIEGRMH NEKVDLMWIC VLYEYLLVGN PPFESASHNE TYRRIVKVDL KFPASVPTGA 300
 QDLISKLRLH NPSEKRLPAQ VSAHPWVRAN SRRVLPSSAL QSWA

Seq ID NO: 204 DNA sequence
 Nucleic Acid Accession #: AK055663
 Coding sequence: 38..1423

1 11 21 31 41 51
 AGAACGGCTT COGGGCGGAG CTGTGCAGCT CCTATCATG GGGACAATTC ATCTCTTTCC 60
 AAAACCACAA AGATCCTTTT TTGGCAAGTT GTTACGGGAA TTTAGACTTG TAGCAGCTGA 120
 CGGAAGGTCC TGGAGATAC TGCTCTTTGG TGTAAATAAC TTGATATGTA CTGGCTTCCT 180
 GCTTATGTGG TCGATTCTTA CTAATAGTAT AGCTTTAACT GCCTATACTT ACCTGACCAT 240
 TTTTGTATCT TTTAGTTTAA TGACATGTTT AATAAGTTAC TGGGTAACTT TGAGGAAACC 300
 TAGCCCTGTC TATTCATTTC GGTITGAAAG ATTAGAAGTC CTGGCTGTAT TTGCCTCCAC 360
 AGTCTTGGCA CAGTGGGAG CTCCTTTTAT ATTAAGAAAG AGTGCAGAAC GCTTTTGGGA 420
 ACAGCCCGAG ATACACACGG GAAGATTATT AGTTGGTACT TTTGGGCTC TTTGTTTCAA 480
 CCGTTTCACG ATGCTTTCTA TTCGSAATAA AOCCTTTGCT TATGTCTCAG AAGCTGCTAG 540
 TAGAGCTGG CTTCAGAGC ATGTTCGAGA TCTTAGTCTA AGCTTGTGTG GAATTATTCC 600
 GGGACTTAGC AGTATCTTCC TTCCCGAAT GAATCCATT GTTTTGATTG ATCTGCTGG 660
 AGCATTTGCT CTTTGTATTA CATATATGCT CATTGAAATT AATAATTATT TTGCCGTAGA 720
 CACTGCCTCT GCTATAGCTA TTGCTTGTAT GACATTGGC ACTATGATC CATTGAGTGT 780
 GTACAGTGGG AAGTCTTAC TCCAGACAAC ACCACCCAT GTTATTGCT AGTTGGACAA 840
 ACTCATCAGA GAGGTATCTA CCTTAGATGG AGTTTTAGAA GTCCGAAATG AACATTTTGT 900
 GACCTAGGT TTTGGCTCAT TGGCTGGATC AGTGCATGTA AGAATTCGAC GAGATGCCAA 960
 TGAACAAATG GTTCTGTCTC ATGTGACCAA CAGGCTGTAC ACTCTAGTGT CTACTCTAAC 1020
 TGTTCAAAT TCAAGGATG ACTGGATTAG GCCTGCTTAA TTGTCTGGGC CTGTTGCAGC 1080
 CAATGTCTTA AACTTTCTAG ATCATCAGT AATCCCAATG CCTCTTTTAA AGGGTACTGA 1140
 TGATTGTAAC CAGTTATCAT CAATCCAGC TAAACCTAGT AGTCCACCTC CAGAAATTTT 1200
 ATTTAACACT CTGGGGAATA ATGGAACCC AGTTATCTT CTAACACAC AAACAAGGCC 1260
 TTATGGTTT GGTCTCAATC ATGGACACAC ACCCTACAGC AGCATGCTTA ATCAAGGACT 1320
 TGGAGTTCCA GGAATTTGGAG CAATCTCAGG ATTGAGGACT GGTTTTACAA ATATACCAAG 1380
 TAGATATGGA ACTAATAATA GAATTGGACA ACCAAGACCA TGATAGACTC TAACTTATT 1440
 TTTAAGGAA TATTGACTCC TTGGCTTCCA ATTTATTTAG TAATCCAACT TTGCATTGAC 1500
 TGTTTAATCA TTTGCTCTAA ATGTTAGATA ATAGTAGTCT TGTTCACATT TCATGAAACC 1560
 TATGAACATA TATTTTGTGA AAATGTATTT GTGACAGTGA AATCCCTGTA AATGTTAAG 1620
 GCTTTAATA GGCTTCCTTT AGAAATGTG TTTCTTAA TTTGGATT TGTATCTTTG 1680
 GTTTGTAGT TGACTGCAGT GTGATGTGAC CTACCTTTA TAAGAGCCAC TTGATGGAGT 1740
 AGATCTGTCA CATTACTAAG ATACGATATT TCTTTTTTT TCCGAGACGG AGTCTGCTC 1800
 TGCCACGTG CCGGACCAAT ACATTATTAT TAACCTAAG CTGACTTTA TTAAGGCTTC 1860
 CTTAGTTTTT GTTTTGTITT GTTTTGTAG ATGGAGTCTC ACTCTGCTGC CAGGCTGGA 1920
 ATGCAGTGGC ATGATCTCAG CTCACCTGAA CCTCTGCTC CTGAGTCAA ATGATCTCC 1980
 TGCCCTCAGC TCCCGAGTAG CTGGGATTAC AGGCACTGAC CACCAAGCCC AGCTAATTT 2040
 TGTATTTTAA GTAAAGACGG GGGATTTTAC CATGTGCGCC AGGCTGCTCT TGAATCTCTG 2100
 ACCCTAGAT CCACCCACCT TAGCTCCCA AAGTGTCTGG ATTAGGTTG AGCCACGCA 2160
 CCTGGCCGAT ATTTTCTTTA ATGGAATTTA TAAATATGCT TCTTGATTA TACACATTT 2220
 GGGAAAGGGA AAAATGTCTG TTTAAAAGT AAAGGTCTCT TTTATAGCT TTTCAAACCT 2280
 AATGTCTAAA TTTTCTTTT AGGTCTCTCT GAATTAATG TACAAACCTA AAAGCAAAA 2340
 TTTTATGAG AAATTTTGA ATACATCTA TCTAGCACA TTTGAATTT TAATTATCAA 2400
 GATTTTTTGT AAAGTTCTC TCCTTAAATA ATTTAGTAT ATTTGTAAT

Seq ID NO: 205 Protein sequence
 Protein Accession #: BAB70980.1

1 11 21 31 41 51
 MGTIHLFRKP QRSFFGKILR EPLVAADRR SWKILLFGVI NLICTGFLLM WCSSTNSIAL 60
 TAYTYLTFED LPSILMTCLIS YNVTLRKPSF VYSFGFERLE VLAVPASTVL AOLGALFILK 120
 ESAERFLEQP EHTHGLLLVG TFVALCFNLF TMLSIRNKPF AYVSEAAST WLQSHVADLS 180
 RSLCGIIPGL SSIFLPRMNE FVLIDLAFAL ALCITYMLIE INNYFAVDTA SRAIALMTF 240
 GTMYEMSVYS GKVLILQTPP HVIGQLDKLI REVSTLDGVL EVRNEHFWTL GFGSLAGSVH 300
 VEIRERDANEQ MVLAVINRL YLVSTLTIV IFKDDMRPA LLGGPVAANV LNFSDHVVIP 360
 MPLLKGTDLL NPTSTPAKP SSPPEFSEFN TPGKNVNPVI LLNTQTRPYG PGLNHGHTFY 420
 SSMLNQGLGV PGIGATQGLR TGFTNIPSKY GTNNRIGQPR P

Seq ID NO: 206 DNA sequence
 Nucleic Acid Accession #: NM_016361.1
 Coding sequence: 397..1662

1 11 21 31 41 51

5
10
15
20
25
30

GGAACTCAGG	GCCGGCTCCT	GTTCCCTTCAA	GAGTGTCTGGA	GGCCAAACTT	GAAATACAAG	60
TTTAATGTTT	CTCGTCGGGC	AAAAGATAAG	GATCCGATCT	CCCCCGGCC	GGTGTGCAGC	120
AGGAGCGACC	AACCCCGACC	CGGGTTAAAA	CTCCACAGGA	CTCTTCGGCTG	CTGCCACCTC	180
TGTTTCTCTC	CCCCGTTCC	ACTCGGGTC	TCCCTCAGGG	CCGGGAGGCA	CAGCGGTCCC	240
TGCTTGTGA	AGGGCTGGAT	GTACGCATCC	GCAGGTTCCC	CGGACTTGG	GGGCGCCCGC	300
TGAGCCCCGG	CGCCCCGAGA	AGACTGTGT	TGCTCTCTG	CAGCTCAAC	CCGGAGGCAG	360
CGAGGGCCTA	CCACCATGAT	CACCTGGTGTG	TTCAGCATGC	GCITGTGGAC	CCCAGTGGGC	420
GTCTTGACCT	CGCTGGCGTA	CTGCCTGCAC	CAGCGGCGGG	TGGCCCTGGC	CGAGCTGCAG	480
GAGGCGGATG	GCCAGTGTCC	GGTCGACCGC	AGCCTGCTGA	AGTTGAAAT	GGTGCAGGTC	540
GTGTTTGGAC	ACGGGGCTCG	GAGTCTCTC	AAGCCGCTCC	CGCTGGAGGA	GCAGGTAGAG	600
TGGAACCCCC	AGCTATTAGA	GGTCCACCC	CAAACTCAGT	TTGATTACAC	AGTCAACCAAT	660
CTAGCTGGTG	GTCCGAAACC	ATATTCTCCT	TACGACTCTC	AATACCATGA	GACCACCCCTG	720
AAGGGGGGCA	TGTTTGTCTG	GCAGCTGACC	AAGGTGGGCA	TGCAGCAAAT	GTTCGCCTTG	780
CGAGAGAGAC	TGAGGAGAA	CTATGTGGAA	GACATTCCCT	TTCTTTCACC	AACCTTCAAC	840
CCACAGGAGG	TCTTTATTCG	TCCACTAAC	ATTTTTCGGA	ATCTGGAGTC	CACCGCTTGT	900
TGCTTGGCTG	GGCTTTTCCA	GTGTGAGAA	GRAGGACCCA	TCATCATCCA	CACGTATGAA	960
GCAGATTGAG	AAGTCTTGTA	TCCCAACTAC	CAAAGCTGCT	GGAGCCTGAG	GCAGAGAACC	1020
AGAGCCCGGA	GGCAGACTGC	CTCTTTACAG	CCAGGAATCT	CAGAGGATTT	GAAAAAGGTG	1080
AAGGACAGGA	TGGGCAATGA	CAGTAGTGAT	AAAGTGGACT	TCTTCATCCT	CCTGGACAAC	1140
GTGGCTGCCG	AGCAGGCACA	CAACCTCCCA	AGCTGCCCCA	TGCTGAAGAG	ATTTGCACGG	1200
ATGATCGAAC	AGAGAGCTGT	GGACACATCC	TGTACATAC	TGCCCAAGGA	AGACAGGGAA	1260
AGTCTTCAGA	TGGCAGTAGG	CCCATTCCCT	CACATCCTAG	AGAGCAACCT	GCTGAAAGCC	1320
ATGGAATCTG	CCACTGCCCC	CGACAGATC	AGAAAGCTGT	ATCTCTATGC	GGCTCATGAT	1380
GTGACCTTCA	TACCGCTCTT	AATGACCCCTG	GGGATTTTTC	ACCAAAATG	GCCACCGTTT	1440
GCTCTGTACC	TGACCATGGA	ACTTTACCAG	CACCTGGAAT	CTAAGGAGTG	GTTTGTGCAG	1500
CTCTATTACC	ACCGGAAGGA	GCAGGTGCCG	AGAGGTTCCT	CTGATGGGCT	CTGCGCGCTG	1560
GACATGTTCT	TGAAATGCCAT	GTGAGTTTAT	ACCTTAAGCC	CAGAAAATA	CCATGCATCT	1620
TGCTCTCAAA	CTCAGGTGAT	GGAAGTTGGA	AATGAAGAGT	AACGTATTTA	TAAAAAGCAG	1680
ATGTTTGAAT	TTTAAATAA	AGTGCCTTTA	TACAAAAAAA	AAAAAATAA	A	

Seq ID NO: 207 Protein sequence
Protein Accession #: NP_057445.1

35
40
45

1	11	21	31	41	51	
NRLMTFVGV	TSLAYCLHQR	RVALAEQEA	DGQCPVDRSL	LKLKMQVVF	RHGARSPLKP	60
LPLEBQVEWN	PQLLEVPQPT	QFDYTVTNLA	GGPKFYSYD	SOYHETTLKG	GMEAGQLTKV	120
GMQMFALGE	RLRKNYVEDI	PFLSPTFNPQ	EVFIRSTNIF	RNLBSTRLCL	AGLFQCCKEG	180
PILIHDEAD	SEVLVFNYS	CNSLRQRTG	RRQTASLQPG	ISEDLLKVKD	RMGIDSSDKV	240
DFPILLDNVA	AEQAHLPLSC	PMLKRPFARM	EQRAVDTSLY	ILPKEDRESL	QMAVGFPLHI	300
LESNLLKAMD	SATADFKIRK	LYLYAHDVT	FIELLMTLGI	FDHKWPPRAV	DLTMELYQHL	360
ESKEWFFVQLY	YHGKQVPRG	CPDGLCPIDM	FLNAMEVYTL	SPERKHEALCS	QTQVMEVGNZ	420
E						

Seq ID NO: 208 DNA sequence
Nucleic Acid Accession #: CAT cluster

50
55
60

1	11	21	31	41	51	
TTTGAGGGGG	TGGTGGGGGG	AGTTTAATTC	ATAAGAAGC	CTCCTGATCA	GAAAGGGGCC	60
TAAACAGCTG	CCCTTGGAGA	GAGTCTCTTC	CTTGAGGATA	AGGCTCCCA	GGGAGGAGG	120
TGCTGGGGGG	CAGTGTTAGG	CTTCAGGCCA	TCCCTGGAGG	CCAGTCCCTGT	GCTCAGCAAG	180
TAGTGGCAGA	GGCTGGAGTG	ATGAGTGGGA	TGGCCTTCTC	AGGTACAGGA	CTGTGCTGCT	240
TCTGGCTGCT	CTTGCAATTC	CAITTGCCAC	TCAGAACTGC	CGCGATCCCA	GCAATGGCCA	300
GGAGCCCTCC	GCAGATCAGT	CGCTCAGCT	GCAGGTTTTT	CCAGTCAATG	TAGAAGGGAT	360
CGTCTTATTT	GCAGAAATGG	TCATTGGCTT	CCAAGGCAGT	CAGGCCAACT	GTGTGACTCT	420
GCAGGTTCTT	CACCTGCTCT	TCACCAAGTG	CCTGCGAGGT	CACCTTGGCG	AGGGCTCAAC	480
TGAGCTGGCA	GCGCAG					

Seq ID NO: 209 DNA sequence
Nucleic Acid Accession #: FGENSEH predicted
Coding sequence: 1..564

65
70
75
80

1	11	21	31	41	51	
ATGGAGCCCT	GGGCGTGGCT	GCAGGGTTTA	AAGAGCCGAC	CCAGTGCCCT	AGCAGCCCTC	60
TCAGATCCGT	TCTCTGCGCT	GCCAGCTCAG	GACACTGGTG	AAGGAGCAGT	GAGGAACCTG	120
CAGAGTCACA	CAGTGGCCCT	GACTGCTTGG	GAAGCCCAATG	ACCCATTTCG	CAATAAAGAC	180
GATCCCTTCT	ACTATGACTG	GAAAAACCTG	CAGCTGAGCG	GACTGATCTG	CGGAGGGCTC	240
CTGGCCATTG	CTGGGATCCG	GGCAGTTCTG	AGTGGCAAAAT	GCAAAATGCA	GAGCAGCCAG	300
AAGCAGACCA	GTCTGTATAC	TGAGAAGGCC	ATCCCACTCA	TCACTCCAGG	CAGATTTCCT	360
ACCTTGGCCA	AATCAATAA	ACCTTTATCT	CCAAGCACCT	TGTCTTGGT	GTTTGGCATC	420
AGCTACACAT	CAGTCTTCCG	AGTGCCCTCT	TCTGCGTCCC	TGTACCTCTG	CATTCTGGT	480
GATGCTGCTG	CCCTACATC	AGGCCATCCA	AGCATGCAGA	ACATAAGCAT	GCAGAACACT	540
GGAACGAAGG	GCTGTACTTA	ATGA				

Seq ID NO: 210 Protein sequence
Protein Accession #: FGENSEH predicted

1	11	21	31	41	51	
MEFWANLQQL	KSRPTCPAAS	SDFFSALPAQ	DTEGAVRNIL	QSHTVGLTAL	EANDPPFANKD	60
DPFYDWNKRL	QLSLGLIOGGL	LALAGIAAVL	SGKCKCKSSQ	KQHSPPVPEKA	IPLITPGRFL	120

TLAKSNKPLS PSTFVLVFGI SYTSVFRVPL SASLYPAIPG DAAALTSQHP SMQNISMONT 180
GTKGCT

Seq ID NO: 211 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..318

1	11	21	31	41	51	
1	11	21	31	41	51	
10	ATGCCCGGCC	ACCCCGTCTG	TGAAGTGAGG	AGCACCTCTG	CCCGGCTGCC	CGCTCTGGGA 60
	AGTGAGGAGC	GCCTCTGCCC	GGCTGCCACC	CGCTCTGTGA	GTGCTGTGCTG	CGCTGGGGCC 120
	AGGCCGCCCC	TGCCCTGCCA	GGCCCTCCGG	CCCCCACCT	TCCACCCACG	GGCTGTCTCC 180
	TCACCCAGG	GTTCCATCTC	CCTAGTTTCC	ACCAGAGACT	GGGTCTTCAT	TCTCACCTCG 240
15	CTACACAGCC	CCTACCAGAA	CGTCTGAAA	TGCAACCTA	ACAACTGTCT	CACCCACGCA 300
	GGAAACTCCC	CAGGCTCCCG	GGCCCCCTGC	GGGGTGTGAG	GCCTCACTCT	TGCGGCCCAT 360
	CCCTCCGCCC	TGACCGCCCT	GAGCTCGCCC	CCAGTGTCTG	CCCTTCAGCT	CCAGTTATCC 420
	CTCCCGCCT	CCAAGGTCCC	CGTTACCGAA	GACCGCCACC	ATCACGACAT	AGCGCAGCAC 480
	ATATGGGACA	CTGGTGAAGG	AGCAGTGAGG	AACCTGCAGA	GTACACAGT	TGGCTGACT 540
20	GCCTTGGGAG	CCAATGACCC	ATTTGCCAAT	AAGAGCGATC	CCCTCTACTA	TGACTGGAAA 600
	AACCTGCAGC	TGAGCGGACT	GATCTGCGGA	GGCTCCTGG	CCATGTCTGG	GATCGCGGCA 660
	GTTCTGAGTG	GCAATGCAA	ATGCAGAGC	AGCCAGAAGC	AGCAGAGTCC	TGTACCTGAG 720
	AAGGCCATCC	CATCTCATAC	TCCAGGCAGA	TTCTCACCT	TGGCCAAATC	AAATAAACCT 780
	TTATCTCCAA	GCACCTTTGT	CTTGGTGTTC	GGCATCAGCT	ACACATCAGT	CTTCCGAGTG 840
25	CTCTTTCTG	CGTCCCTGTA	CCCTGCCATT	CCTGGTGAATG	CTGCTGCCCT	CACATCAGGC 900
	CATCCAAGCA	TGCAGAACAT	AAGCATGCAG	AACACTGGAA	CGAAGGCTG	TACCTAA

Seq ID NO: 212 Protein sequence
Protein Accession #: FGENESH predicted

1	11	21	31	41	51	
1	11	21	31	41	51	
30	MPGHPVCEVR	STEARLPRLG	SEERLCPAAT	PSVSACCAGP	RPPVPQALR	PPTFHFRACS 60
	SPQGSISLVS	TRDWVFILTL	LHSPYQNVLK	CKPNNCLTPA	GNSPGRAPC	GVAGLTLRAH 120
35	PSALTALSSP	PVLALHVLQS	LPASKVEVTE	DRHHEDIQOH	IWDTEGAVR	NLQSHTVGLT 180
	ALEANDPPAN	KDDPFYVDMK	NLQLSGLICG	GLLAITAGIAA	VLSGCKCKCS	SQKQHSFPVPE 240
	KAIPILITGR	FMTLAKSNKP	LSPSTFVLVF	GISYTSVFRV	PLSASLYPAI	PGDAAALTSQ 300
	HPSMQNISMQ	NTGTGKCT				

Seq ID NO: 213 DNA sequence
Nucleic Acid Accession #: FGENESH predicted
Coding sequence: 1..1758

1	11	21	31	41	51	
1	11	21	31	41	51	
45	ATGATGGGGT	CTCATGTTCG	CCAGGCTGGT	CITGAACCTCC	TGGGCTCGAG	TGACCCCTCCT 60
	GCCTTGGCT	CCGAAAGTGC	TGGGATTACA	GGACTGTAT	TACAGGAATC	CATAACACTG 120
	GAGGATGTGG	CTGTGGACTT	CACTTGGGAG	GAGTGGCAAC	TCTTGGGCGC	TGCTCAGAG 180
	GACCTGTACC	GGGATGTGAT	GTGAGAGAAC	TACAGCAACC	TGGTGGCAGT	GGGGTATCAA 240
50	GCCAGCAAA	CGGATGCAC	CTTCAAGTTG	GAACAAGGAG	AACAACCTGT	GACAATTGAA 300
	GATGGAAATCC	ACAGTGGAGC	CTGTTCAGGT	TCTCCAAAGG	TCCCGTTCCT	CATTTTCTCA 360
	TCTGTGCCCT	TCATCTCTCA	AAATTGCTT	CATTCTAACA	TATGGAAAGT	TGATCATGTG 420
	CTGGAGGCGT	TGCAGAGTGA	AAGCCTGGTG	AACAGAAAGG	AACCATGTCA	TGAACATGAT 480
	GCATTGAAA	ATATTTGTTC	TTGCAGCAAA	AGTCAGTTTC	TGTTAGGGCA	AAATCATGAT 540
55	ATATTTGACT	TACGTGAGAA	AAGTTTGAAA	TCCATTTTAA	CTTTAGTTAA	CCAGAGCAAA 600
	GGCTATGAAA	TAAAGAACTC	TGTTAGTTTC	ACTGGAAATG	GGGACTTCCT	TCTTCATGCT 660
	AACCATGAAC	GACTTCATAC	TGCAATTAAA	TTCCCTGCAA	GTCAAAACT	CATCAGCACT 720
	AAGTCCCAAT	TCATCAGTCC	CAAGCATCAG	AAAACAGGAA	AAATTAGAGAA	GCATCATGTG 780
	TGCAGTGAAT	TGGGGAAGGC	CTTCATCAAG	AAGTCTTGCC	TAACTGATCA	CCAGGTAATG 840
60	CATACAGAG	AGAAACCCCA	CAGATGTAGT	CTATGTGAGA	AAGCCTTCTC	CAGAAAGTTC 900
	ATGCTTACTG	AACATCAGCG	AACATCATCA	GGAGAAAAC	CTTATGAATG	CCCTGAATGT 960
	GGCAAGGCT	TTCTCAGAAA	ATCACGGCTC	AACATACATC	AGAAAACACA	TACCGGAGAG 1020
	AAACCTTATA	TATGCACTGA	ATGTGGAAAA	GGCTTCATCC	AGAAAGGAAA	TCTCATTTGA 1080
	CACCGGOGAA	TTCATACAGG	TGAGAAACCT	TATATATGCA	ATGAATGTGG	AAAAGGCTTC 1140
65	ATTCAGAGAA	CGTGTCTCAT	AGCACATCAG	AGATTTCACA	CAGGAAGAGC	GGCCTTTGTG 1200
	TGCAGTGAAT	GTGGAAAATC	CTGTTCTCAG	AAATCAGGTC	TCATTAAACA	TCAAAGAAIT 1260
	CACACAGGAG	AGAAACCCCT	TGAATGTAGT	GAAATGTGGA	AAGCCTTTAG	CACAAAGCAA 1320
	AAGCTCATTT	TCCATCAAGG	GACTCATACA	GGAGAGAGAC	CCTATGGCTG	TAAAGAGTGT 1380
	GGCAAGCGT	TTGCGTATAT	GTGTTGCTG	GTAAAGCATA	AGAGATATCA	CACAAAGGAG 1440
70	AAACAAGAGG	CAGCCAGAGT	GGAAAATCCT	CCTGCAGAGA	GGCAGAGCTC	ATTACACACC 1500
	AGTGATGTCA	TGCAGGAGAA	AAACTCTGCT	AACGGGGGGA	CTACACAGT	GCCTTCTGTG 1560
	GCCCTCAGA	CATCATTAATA	CATCAGCGGC	CTCCTGGCAA	ACAGGAACGT	AGTCTCTGTG 1620
	GGACAGCCAG	TGGTCAAGAT	TGCAGCCTCA	GGAGATACCA	GAGGATTTGC	ACAGGACAGA 1680
75	AACCTTGTA	ATGCAGTGAA	TGTGGTTGTG	CCTCCGTGA	TCAATTATGT	CTTATTTTAT 1740
	GTTACAGAAA	ACCCATAG				

Seq ID NO: 214 Protein sequence
Protein Accession #: FGENESH predicted

1	11	21	31	41	51	
1	11	21	31	41	51	
80	MNGSEVAQAG	LELLGSSDPP	ALASESAGIT	GLLLQESITL	EDVAVDFTWE	EWQLLGAAQK 60
	DLYRDMLEN	YSLNVAVGVD	ASKFDALFKL	BQGBQLWTIE	DGHSAGACBG	SPKVPPSIPS 120
	SVPFTLQACL	HSNWKVDHV	LERLQESLIV	NRRKPCHEED	AFENIVHCCK	SQFLGQNEED 180
	IFDLRGSLK	SNITLVNQSK	GYEIKNSVEF	TGNGDSFLHA	NHERLITATK	FPASQKLIST 240

KSQFISPIHQ KTRKLEKHHV CSECGKAFIK KSWLTDHQM HTGKPHRCS LCEKAFSRKF 300
 MLTEHQRTHT GEKPYECPEC GKAFLLKSLR NIHQRTHTGE KPYICSECGK GFIOKGNLIV 360
 HQRIHTGEKP YICNECKGKF IQRTCLIAHQ RPHTGKTPFV CSECGKSCSQ KSLGLIKHQRI 420
 HTGKPPFECB ECGKAFSTKQ KLIVHQRTHT GERPYGCNEC GKAFAYMSCL VKHKRIHTRE 480
 KQEAARKVENP PAERHSSLHT SDVMQENSA NGATTQVPSV APQTSIATISG LILANRNVVLV 540
 GQPVVRCAAS GDNRRGPAQDR NLVNAVNVVV PSVINYLIFY VTEHP

Seq ID NO: 215 DNA sequence
 Nucleic Acid Accession #: NM_032190.1
 Coding sequence: 502..1332

1 11 21 31 41 51
 | | | | | |
 GATTCGGTGT TCTTGGCCAT GTTAGCCATA ATATCCTGTG CAGTATGTTT TCCTGTGCA 60
 GAGGCAAAAA CATATTGGGC ATATGTTCCC AAGCCGCCAG CAGTATGACC CATACTTTGG 120
 AGTGACACTC CTCTAAGAT TTATCATGAT TAAGGAGCAT GGGCTCCAGG ACCCTTAACF 180
 CCACCTGACA TAGAACAGTT AGACTCTCAG AATAATGTCA TTAATTATAC CGCTCCATTG 240
 GAAGGACTTC CTTTGTGTGT CACCACAAAG ACATCACTCA GCCATAGCTG TCTTACAGTT 300
 CAAGCTCACA CATGGTTGAG TCACATATGGG AAAATCATGT ACTTATTAAAG TCTTGGTTAT 360
 ATTAATGTAA CCGGTGTGCT AACCAACCAT TCCTGGCCCA ATCGCCTTCA TTGTGCTGAC 420
 TATACAGATG GGATTCCTCT CAATAGTTCC TACCCCTCTC CATAGACCCA GTGTCTTGGC 480
 CCACCTGCTA GAAACAAATC TATGTTAACT GGAGACATTG TGGATTGGGG ACCTAAAGGC 540
 CAATTAGATG GAAAGAGAGA AAATCAGAAA TCGTGGCACA AACTTGTCTG GCATTGGTGG 600
 CAGGCTTTTA ATGCTTCTTC TTTTATAAAC ACTGGGATCC AATCCAGATC GGCCTGCCAG 660
 ATTGCTTGGC ATGGAGCAGG CTTTAGCCCG CCTCTTCTTC AGTGGCAITTA TCTAGGGAGG 720
 AAGAGACCAA TTCAAAAGAT GATATGGAGG GCAGCATTCG CATTATGAA TGGCAACATC 780
 TGGGTGGCCA TAATACATC CAATAATAGC AATAGTAGGC AACACAGTCT TAATGTTACA 840
 TTTGTAAAGA ATATCACCAC TCAATTTACA GTTGTGTTT TTAATCCTTA TGTGTTTTTG 900
 GCAGCTAAGA AGGACCAAGT CCAGGTAAAC AATACCCAAT TGACCTGTAA ATCTTGGCAG 960
 TTATATCACT GATTAATACA TAGCACATGG CAAACACATA ATATCTCTAC TTTGATGATT 1020
 TTAGGTGCA TCCCTGGGCT ATGGATTCTT GTTAATCTGT CTGAGCCATG GGTGCTCCACA 1080
 ATTGCTTTAC AITTTGTGAA ACTTCTTCTA ACTCAGTTTA CTCATTGTGT CCGTAGAGGC 1140
 TTAGGCATGA TAATTTTTCG TATTGTTTAC TTGGTCCAC TAATAATTTT TGTGTGTGATG 1200
 TCCTCTGTAG CTTTGCATAG TTCTATTCAA ACAGCTCAGT ATGTGGAGAA CTGGACACGC 1260
 ACAGTCAACC AAGGTGTGCT ACTTGAGAA AAAATTAACA CTGAGTTACA AACTGAAGTG 1320
 GCAGTGTAT AATCCACGAT TCTATGGTTA GGGGAACAG TACAAGCTTT GCAATTGCAG 1380
 CAGTAATGT GTGTCTCAT TATCACACT CATATTGTG TAAACCACTT AGAATATAAC 1440
 CAAAGTGAGT ATCCATGGGA TCTTGTGAAA GCCCATTTGC AGGGAGCTTT CACATCCGAC 1500
 ATCACTTTTG ATATTGGTGA ATTACAAAC AAAATTCTTG ATTTAAATAA ACAAATTTCA 1560
 GAGTTTCAGC CTTCTTTAGA AGACTGGACT GAATCCAGC AAGGCTTGGA GAGGCTCAAC 1620
 CCTTGGACCT ATCTAAAGCA CCACATTAAC ATCTATATA TAGTTCTTGG AATAATGTTG 1680
 TTTGTCTCT GTCTCTGTG TATAGTCTGT AAAATCGGAT GGACTGCCAA TCGGAGATG 1740
 AAAGCTACCC AGCCTGGCCT TACATCTTT CACTTAATAC ATAAACAGA AGGGGGAANT 1800
 GTTGGGAGCC AAAAGGCCA AAGGGATGAT GACCAACTCA GCATTCCACT GGAGGCTCAC 1860
 TGATCAAAAC GCAAACTGTT TATCATGAAT ACAGAAATGT GGCAAACTCG CTTCTGTGCC 1920
 TGCCAGAGAG GTTGTCTGAG GGCCATGCT CCTTGGCCCC GGCTCCTTGA GGTATCTTAC 1980
 TGGGACATCT AGAGCCTATT GTTCGAGGAA TGCAGTCTG CAAGCCTACT CTGGACGAG 2040
 CAGCTGACCT CTTCTTCCAC ACCCTTCTCT ACTATCTCT TTGCCTAATA AATATGGAGG 2100
 GCTGTGTAAA GCTCAGGGCC CTTGTCCACT AGAGGCAAGG TGTCCCTTGA CCTTCTTCC 2160
 AACAT

Seq ID NO: 216 Protein sequence
 Protein Accession #: NP_115566.1

1 11 21 31 41 51
 | | | | | |
 MLTGDIVDNG PRGOLDGKEE NQKSWHKLW HWWQAFNASS LYNTGIQSQS AAQIANEBGAG 60
 FSPPLPQWHY LGRKGPQKM IWKAAPFPMN GNIWVAIILS MNSNSKQHSI NVIFVKNTIT 120
 QFTVCVFNFY VFLAAKKDQL QVNTQLTCK SQQLYHCINH STLTQHMIST LMLGCIPLGL 180
 WIPVLSSEFW AATIALRPFVK LLLTQPTHCV RRLGLMIIFA IVYLVTLIIS VVMSVVALRS 240
 SIQTAQVYEN WTRTVNQGL LENKINTELQ TEVAVL

Seq ID NO: 217 DNA sequence
 Nucleic Acid Accession #: FGENESH predicted
 Coding sequence: 1..1566

1 11 21 31 41 51
 | | | | | |
 ATGGTGAACC CCAATCCAC TTCTCTCTC TTCAGTTAT GTTTTGTCT CCTGAGGAGT 60
 CAGAACCTGT GGGTTGAGGA GCAAAATCAA TGCAAAACA TATTGGGCAT ATGTTCCCAA 120
 TCCCCAGACA GTATGGCCTA TACTTTGGAG CTCACCTCTC CTGAGATTTA TCAAGATCAG 180
 GGAGAGTGGG CTCACAGGAC CCTAATCCC CGTGACATAG AAAAGTTAGA CTTCTCAGAAC 240
 AATGTCATTA ATTATACCAC TCCACTGAAA GGACTCCCTT TGTTTATCAC CACAAAGAGG 300
 TCGCTCAGCC ATAGCTGTCT TGCAATTCAA GCTCRAACAT GGTGTAGTCA CTATGGAAAA 360
 ATTATGTACT TATTAGGTCT TGGTTCTATT AATGTAACTG GTGTGCTAAC CAATCATTCG 420
 CAGTCCAGTC ACCCTAATTTG TGCTGATTAT ACAGAAATGA TTCCATTCAA TAGTTOCTAC 480
 CCCACTCTGT GGACCCAGTG TCTGTATCCA CTGGCTAGTA AACAAATATAT GTCAACTGAA 540
 GACACTGTGG ATTGGGAACC TAAAGGTCAA TTAGATGGAA AAGGTGAAG TCAGAAATCA 600
 TGGCAAAAC TCCACTGGCA TTGGCGGCAA GCCTTTAATG CTTCTCTTIT ATACAACAGC 660
 AGAATCAAACT CCCAGTCTGC TGCTCAGATT GCCTTGCATG GAGCAGGCTT TAGCCCACTT 720
 CTCTCTCAGT TGCATTATCT GGGGAGGAAA GGACCAATTC AAGAAACTAT ATGGAAGGCA 780
 GCACTCCCAT TATGAAATGG CAACATCTGG ATTGGAACAC TGTCTAATAA TAGCAATAGT 840
 AAGCAACACA GTCTTAATGT TGCAATTGTA AAGAAATACA CCACCTCAGT TACAGTTTGT 900
 GTTTTAAATC CTTATGCCCT TTTGGCAGCT AAGAGAACC AGCTTCAGGT GGAGAACTGG 960

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

```

ACACGCACAG CTGACCAAGC GAGGCTACTT CAGAATAAAA TTAACACTGA GTTACAAACT 1020
GAAGTGGCAA TGTGAAATC CATGGTTCTG TGGTTAGGAG AACAGGTACA AAGCTTGCG 1080
TTGCAGCAGC AATTGCGTCA TCATTTTAAT CACATTCATA TTTGCGTAAC TAACTCAGAA 1140
TATAACCAAA GTGAGTATCC GTGCGACCTT GTGAAAGCCC ATTTGCAAGG AGCTTTCACA 1200
TCCAACATCA CCTTTGATAT TGGTGAATTA CAAAACAAAA TTATTGATTT AAATAGGCAA 1260
ACTCAAGBAAT TTCAGCCTTC TTTAGAAGAC TGGACCGAAT TCCAGGAAGG CCTGGAGAGC 1320
CTCAACCCCT GGACCTATCT AAGGCACCAC ATTAACATCT TATATGTAGT TCTTGAATA 1380
ATGTTGTTTT GTCTCTGTCT TCGGTTTATA GTCTGTAAAA TCGGATGGAC CACCAATTGG 1440
AGAATGAGAG CCTCCAGGCC CAGCCTTACA TTCTTCAAT TAATACATAA ACAGAAAGGG 1500
GGATATGCGG GGAGCCAAAG GCCTGTGGGA CGTGACCAAC TCAGCATTCT GCTGGAGGCT 1560
ATATGA

```

Seq ID NO: 218 Protein sequence

Protein Accession #: FGENESH predicted

1 11 21 31 41 51

```

MVNPKSTSSL FRLCFLLLRS QNLWVERQIQ CKNILGICSD SPSSMAYTLE LTPPETIYHDQ 60
GEWAPGGLTP RDIEKLDSON NVINYTTPLB GLPLFITTKT SLSSHSCLAIQ AQTWLSHYGK 120
IMVLLGLGSI NVITGLTNHS QSSHPNCADY TEWTFPNSY PTLMTQCCLDP LASKQYMSYE 180
DTVDWEKKGQ LDGKGESQKS WHKLHWHNRQ AENASSLYNS RIQSGSAAQI ANHAGGFSFP 240
LPQLHYLGRK GPIQSTIWK AALPFMGNWNI IGTLSNNSNS KQHSINVAFV KNITTOFTVC 300
VFNPYAFIAA KKNQLQVENW TKTADQARLL QNKINTEIQT EVAMLKSMVL WLGSQVQSLQ 360
LQQQLRHHFN HIHICVTNBS YNQSEYFWDL VKAHLQGAFT SNITFDIGEL QNKIIDLNRO 420
TQEFQPSLE WTEFQESLES LNPWTYLRSH INILYVVLGI MLFCLCLRFI VCKIGWTTNW 480
RMRASQPSLT PFQLIHKQKG GYAGSQRPVG RDQLSILLEA I

```

Seq ID NO: 219 DNA sequence

Nucleic Acid Accession #: FGENESH predicted

Coding sequence: 1..900

1 11 21 31 41 51

```

ATGCCGCGCG GGGAGCTGAG CGAGGCGGAG CCGCCCCCGC TCGCGGCCCC GACCCCTCCC 60
CCGCGGCGGC GTAGCGCGCC CCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC 120
GCGGTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180
CGGCCCCACTG CGCTGGACAC CTCTCTCTGT ACCTACGTTT AATCGCCCGT GGGCGCGCGT 240
GGCTGCGGCG GGGCTGTGCA CCGGGGAGCT GGGGCGGGCG TCTCGGCGGG AGGGCGCAGA 300
GGACCCCGGG GAGGAGACTG GAGCAGGCCC CGAGGTGGGG CTGGTGGCGC CCAGGACGCT 360
CTTCTTAATC CAGGCTCTCC CCGCCCGGCC CTGCGAGTGC AAGTCTGGT GAGTGGAGCT 420
CCGCTGCGCA TTGAGCTCTG GGACACAGCG GGACAGGAGG ATTTTGACCG ACTTCGTTC 480
CTTTGCTACC CGGATACGGA TGCTCTCTG GCGTGCTTCA GCGTGGTGA GCGCCAGCTC 540
TTTCAAAACA TCACAGAGAA ATGGCTGCCC GAGATCCGCA CGCACAAACC CCAGGCGCCT 600
GTGCTGCTGG TGGGCACCCA GCGCGACCTG AGGGACGATG TCAACGTACT AATTGAGCTG 660
GACCAAGGGG GCGGGGAGGG CCGCGTGGCC CAACCCAGG CTCAGGGTCT GGCCGAGAAG 720
ATCCGAGCTG GCTGCTACCT TGAAGTCTCA GCCTTGACG AGAAGAACTT GAAGGAAGTA 780
TTTGACTCGG CTATTCCTAG TGCATTGAG CACAAAGCCC GCGTGGAGAA GAAACTGAAT 840
GCCAAGGTG TCGCACCCCT CTCCCGCTGC CGCTGGAAGA AGTTCTCTG CTTCGTTTGA

```

Seq ID NO: 220 Protein sequence

Protein Accession #: FGENESH predicted

1 11 21 31 41 51

```

MFFRELSEAE PPFLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIVS YTCNGVPARY 60
RPTALDTPFG TYVQSPVFRF GCGGAVHRGA GAGVBAGRR GPRGGDWSRP EGGAGAAQDA 120
LNSGSPRFA PAVQVLVDGA FVRIELNDTA GQEDFRLRS LCYEDTDVFL ACFSVVQPS 180
FQNIYKWLFP BIRTHNPQAF VLLVGTQADL RDDVNVLIQL DQGGREGFVP QPQAQGLAEK 240
IRACYLECS ALTQKNLKEV FDSAILSABE HKARLEKKLN AKGVRTLSRC RWKKFFCFV

```

Seq ID NO: 221 DNA sequence

Nucleic Acid Accession #: XM_063832.2

Coding sequence: 1..711

1 11 21 31 41 51

```

ATGCCGCGCG GGGAGCTGAG CGAGGCGGAG CCGCCCCCGC TCGCGGCCCC GACCCCTCCC 60
CCGCGGCGGC GTAGCGCGCC CCAGAGCTG GGCATCAAGT GCGTGCTGGT GGGCGACGGC 120
GCGGTGGGCA AGAGCAGCCT CATCGTCAGC TACACCTGCA ATGGGTACCC CGCGCGCTAC 180
CGGCCCCACTG CGCTGGACAC CTCTCTCTGT CAAGTCTCTG TGGATGGAGC TCGGCTGCG 240
ATTGAGCTCT GGGACACAGG GGGACAGGAG GATTTTGACC GACTTCGTTC CCTTTGCTAC 300
CGGATACGAG ATGTCTTCTT GCGTGCTTTC AGCGTGGTGC AGCCCGAGCTC CTTTCAAAAC 360
ATCACAGAGA AATGGCTGCT CGAGATCCGC ACGCACAAAC CCCAGGCGCC TGTGCTGCTG 420
GTGGGCACCC AGGCGGACCT GAGGAGGAT GTCAACGTAC TAATTCAGCT GGACCAAGGG 480
GGCCGGGAGG GCGCGGTGCC CCAACCCAG GCTCAGGGTC TGGCGGAGAA GATCCGAGCC 540
TGCTGCTACC TTGAGTGTCT AGCCTTGACG CAGAAGAACT TGAAGGAAGT ATTTGACTCG 600
GCTATTCCTA GTGCCATTGA GCACAAAGCC CGGCTGGAGA AGAAACTGAA TGCCAAAGGT 660
GTGCGCAGCC TCTCCGCTG CCGCTGGAAG AAGTTCTTCT GCTTGTGTTG A

```

Seq ID NO: 222 Protein sequence

Protein Accession #: XP_063832.1

1 11 21 31 41 51

MPPRELSEAE PPFLRAPTPP PRRRSAPPEL GIKCVLVGDG AVGKSSLIYS YTCNGYPARY 60
 RPTALDIFSV QVLVDGAPVR IELNDTAGOE DFDRLRSLCY PDTDVFLACF SVVQSSSPON 120
 ITERKWLPEIR TBNPQAPVLL VGTQADLRDD VNVLIQLDQC GREBPVPQPP AQGLAEKIRA 180
 CCYLECSALT OKNLKEVFDG AILSAIEHKA ELEKGLNAKG VRTLSRCRWK KPFCFV

Seq ID NO: 223 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..1161

1 11 21 31 41 51
 ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 60
 TTOCGAGATG ACTTCATTGC CAAGGTGTGG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT 120
 GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTCTGTGT TCCACCTCAA GTCTGGGAAA 180
 TCCAGCCGGA TTFTCTGTGT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 240
 CCGTCTCGTA TGGACTACTA TGTGCGCGGT TCAGACTGGA AGTTTGGGGA CATCCCTTGC 300
 CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCAGGTG 360
 GTGGCGGTAG ACAGGTATTT CCGGTGTGTC CATCCCCACC ACGCCCTGAA CAGATCTCC 420
 AATFGGACAG CAGCCATCAT CTCTTGCCCT CTGTGGGGCA TCACTGTGTG CTAACAGTC 480
 CACCTCTGTA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC 540
 AGCATCTGCC ATACCTTCOG GTGGCAGGAA GCTATGTTCC TCTGTGAGTT CTTCTGCCC 600
 CTGGGCATCA TCTCTGTTCT CTCAGCCAGA ATTATCTGGA GCTGCGGCA GAGACAAATG 660
 GACCGGCATG CCAAGATCAA GAGAGCCATC ACCCTCATCA TGGTGTGTGG CATGCTCTTT 720
 GTCATCTGCT TCCTTCCAG CGTGGTGTGG CCGATCCGCA TCTTCTGGCT CTTGCACTCT 780
 TOGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCTTCTT TATCACTCTC 840
 AGCTTCACCT ACATGAACAG CATGCTGGAC CCGCTGTGTG ACTACTTCTC CAGCCCATCC 900
 TTTCOCRACT TCTTCTCCAC TTTGATCAAC CGCTGCTTCC AGAGGAAGAT GACAGGTGAG 960
 CCAGATAATA ACCGACGAC GAGCGTCSAG CTCACAGGGG ACCCCAACAA AACGAGAGGC 1020
 GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCTCTTA TCTGGGCCCA 1080
 ACCTCAATA ACCATTCCAA GAAGGACAT TGTCAACAAG AACCAGCATC TCTGGAGAAA 1140
 CAGTTGGGAT GTTGCATCGA G

Seq ID NO: 224 Protein sequence
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 MNRHQLQDEF LEIDKKNCCV FRDDPIAKVL PPVLGLEFIF GLLNGLALW IPCFHLKSWK 60
 SSRIFLNLVA VADFLLIICL PFMVDYVRR SDWKFGDIPC RLVLFMFAMN EQGSIIPLTV 120
 VAVDRYFRVV HPHHALNKLIS NRTAAIISCL LMGITVGLTV HLLKKKLIQ NGPANVCISF 180
 SICHTFQWHE AMFLLEFLFP LGIILFCSAR IINSLRQRM DRHAKIKRAI TFMVVAIVF 240
 VICLEPSVVV RIRIFWLLHT SGTQNCVYR SVDLAFFITL SFTYMNMLD PVVYFSSPS 300
 PFNFFSTLIN RCLQRKMTGE PDNNRSTSVL LTGDPNKTRG APEALMANSQ EPWSPSYLGP 360
 TSNHNSKKGH CHQEPASLEK QLGCCTE

Seq ID NO: 225 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 1..1092

1 11 21 31 41 51
 ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 60
 TTOCGAGATG ACTTCATTGT CAAGGTGTGG CCGCCGGTGT TGGGGCTGGA GTTTATCTTC 120
 GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTCTGTGT TCCACCTCAA GTCTGGGAAA 180
 TCCAGCCGGA TTFTCTGTGT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 240
 CCGTCTCGTA TGGACTACTA TGTGAGGCGT TGGGACTGGA AGTTTGGGGA CATCCCTTGC 300
 CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCAGGTG 360
 GTGGCGGTAG ACAGGTATTT CCGGTGTGTC CATCCCCACC ACGCCCTGAA CAGATCTCC 420
 AATCGACAG CAGCCATCAT CTCTTGCCCT CTGTGGGGCA TCACTATGCG CTTGACAGTC 480
 CACCTCTGTA AGAAGAAGAT GCGATCCAG AATGGCGGTG CAAATTTGTG CAGCAGCTTC 540
 AGCATCTGCC ATACCTTCCA GTGGCAGGAA GCTATGTTCC TCTGTGAGTT CTTCTGCCC 600
 CTGGGCATCA TCTCTGTTCT CTCAGCCAGA ATTATCTGGA GCTGCGGCA GAGACAAATG 660
 GACCGGCATG CCAAGATCAA GAGAGCCATC ACCCTCATCA TGGTGTGTGG CATGCTCTTT 720
 GTCATCTGCT TCCTTCCAG CGTGGTGTGG CCGATCCGCA TCTTCTGGCT CTTGCACTCT 780
 TOGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCTTCTT TATCACTCTC 840
 AGCTTCACCT ACATGAACAG CATGCTGGAC CCGCTGTGTG ACTACTTCTC CAGCCCATCC 900
 TTTCOCRACT TCTTCTCCAC TTTGATCAAC CGCTGCTTCC AGAGGAAGAT GACAGGTGAG 960
 CCAGATAATA ACCGACGAC GAGCGTCSAG CTCACAGGGG ACCCCAACAA AACGAGAGGC 1020
 GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCTCTTA TCTGGGCCCA 1080

Seq ID NO: 226 Protein sequence
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 MNRHQLQDEF LEIDKKNCCV FRDDFIVKVL PPVLGLEFIF GLLNGLALW IPCFHLKSWK 60
 SSRIFLNLVA VADFLLIICL PFLMDYVRR WDWKPGDIPC RLVLFMFAMN EQGSIIPLTV 120
 VAVDRYFRVV HPHHALNKLIS NRTAAIISCL LMGITVGLTV HLLKKKLIQ NGGANLCSF 180
 SICHTFQWHE AMFLLEFLFP LGIILFCSAR IINSLRQRM DRHAKIKRAI TFMVVAIVF 240
 VICLEPSVVV RIRIFWLLHT SGTQNCVYR SVDLAFFITL SFTYMNMLD PVVYFSSPS 300
 PFNFFSTLIN RCLQRKMTGE PDNNRSTSVL LTGDPNKTRG APEALMANSQ EPWSPSYLGP 360
 TSP

Seq ID NO: 227 DNA sequence
Nucleic Acid Accession #: NM_006018
Coding sequence: 61..1224

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
      CGCCACTTGG CTGGAGCAATT CACTAGGCGA GCGGCTCCAT CGGACTCACT AGCCGCACTC 60
      ATGAATGGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 120
10     TTCCGAGATG ACTTCATTGC CAAGGTGTGT CCGCCGCTGT TGGGCTGGA GTTTATCTTT 180
      GGGCTTCTGG GCATATGGCT TGCCCTGTGG ATTTCTGTGT TCCACTCAAA GTCTGGAAA 240
      TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 300
      CCGTTGTGTA TGGACTACTA TGTGCGGCGT TCAGACTGGA ACTTTGGGGA CATCCCTTGC 360
      CGGCTGTGTC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG 420
      GTGGCGGTAG ACAGGTATTT CCGGCTGGTC CATCCCAACC ACCGCTTGA CAAGATCTCC 480
15     AATGGACAG CAGCCATCAT CTCTGCTT CTGTGGGGA TCACGTGTGG CCTAACAGTC 540
      CACCTCTCTA AGAGAGAGTT GCTGATCCAG AATGGCCCTG CAATGTGTGT CATCAGCTTC 600
      AGCATCTGCC ATACTTCCG GTGGCAGGAA GCTATGTTCC TCCTGGAGTT CCTCCTGCC 660
      CTGGGCATCA TCCGTTCCTG CTGAGCCAGA ATTATCTGGA GCTGCGGCA GAGACAAATG 720
      GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGTGGC CATCGTCTTT 780
20     GTCACTGTCT TCTTCCCGAG CGTGTGTGT GCGATCCGCA TCTTCTGGCT CTGCACACT 840
      TCGGGCAGTC AGAATGTGTA AGTGTACCGC TCGTGGACC TGGCTTCTT TATCACTCTC 900
      AGCTTCACCT ACATGAACAG CATGCTGGAC CCGTGGTGT ACTACTTCTC CAGCCCATCC 960
      TTTCCCAACT TCTTCTCCAC TTGTATCAAC CGCTGCTCC AGAGGAAGAT GACAGGTGAG 1020
      CCAGATAATA ACOCGAGCAC GAGCGTGGAG CTCACAGGGG ACCCCCAACA AACAGAGGGC 1080
25     GCTCCAGAGC CATTATATGG CACTCCGCT GAGCCATGGA GCGCTCTTA TCTGGGCCA 1140
      ACCTCAATAA ACCATTCCAA GAAGGGACAT TGTACCAAG AACCAATC TCTGGAGAAA 1200
      CAGTTGGGCT GTTGCATCGA GTAATGTAC TGGACTCGGC CTAAGGTTTC CTGGAACCTC 1260
      CAGATTCAAG GAATCTGATT TAGGAAACT GTGGCAGATG AGTGGGAGAG TGGTTGCAAG 1320
      GTGTGACCA AGGAATCTGT GAGGAACAGA GATTAAGCT TCTAGGCATC TGAACCTTGC 1380
30     TTCACTCTG ACGTCCGAG GACTGAAGAT GGCCTAATTG TAGGCGTTTC TGCTGAGCAG 1440
      AGTTGAGGCC AGAGATCTAC TTGTGACTTG TTGGCTCTCT TCCACATCT GCCTCAGACT 1500
      GGGGGGGGCT CAGCTCTCTG GGTGATATCT AGCCTGCTTG TGAGCTCTAG CAGGGATAAG 1560
      GAGAGCTGAG ATTGGAGGGA ATTGTGTTGC TCCTGGAGGA AGCCCAAGCA TCATTAAACA 1620
      AGCCAGTAGG TCACCTGGCT TCCGTGGACC AATTCTATCT TCAGACAAGC TTTAGAGAAA 1680
35     TGGACTCAG GAAGAGACTC ACATGCTTTG GTTAGTATCT GTGTTTCCGG TGGGTGTAAT 1740
      AGGGGATTAG CCCCAGAGCG GACTGAGCTA AACAGTGTTA TTATGGGAAA GGAATAGGCA 1800
      TTGCTGCTTT CAAACAGCGA CTAATGCAAT CCATCTCTCT CTGTGTTATA GTAATCTAAG 1860
      GGTGAGCTAG TTAAGACGCG TTCAGGATAG AAGCTGTGTT CCCACTGTGT TCGTTTAC 1920
40     ATTAAGAGG AAGCTGCTCT CTGCCCCACG GGTAGAGGGG GTGCACGTTT CTCTGTGTTT 1980
      CTTGCTGTGT GTTCTGTATC TTACCAAAA TCTACCATT CAATAAATT TGATAGGAGA 2040
      CAAAAAAA A

```

Seq ID NO: 228 Protein sequence
Protein Accession #: NP_006009.1

```

45     1      11      21      31      41      51
      |      |      |      |      |      |
      MNRHHLQDHF LEIDKKNCCV FRDDFIKVL PVVLGLEPFI GLLEGLALW IFCFHLKSWK 60
      SSRIFLENLA VADFLLIICL PFVMDYVYRR SDWNFQDIPC RLVLPMFAMN RQGSIIPLTV 120
50     VADDEYFRV HPHEALNCKS NWTAAIISCL LWGITVGLTV HLLKKLLIQ NGPANVCISF 180
      SICHTFWRHE AMELLEFLLE LGIILFCSAR IINSLRGRQM DRHAKIKRAI TFMVVAIVF 240
      VICFLPSVVV RIRIFWLLHT SGTONCEVYR SVDLAFITL SPTMNSMLD PVVYFSSPS 300
      EPNFFSTLIN RCLQRMTGSE PDNRRSTEVE LTGDPNKTG APEALMANSG BPWSPSYLGP 360
55     TSNHNSKKGH CQRPASLEK QLGCCIE

```

Seq ID NO: 229 DNA sequence
Nucleic Acid Accession #: NM_014398.1
Coding sequence: 64..1314

```

60     1      11      21      31      41      51
      |      |      |      |      |      |
      GGCACCGATT CCGGCGCTGC CCGGACTTCG CCGCACGCTG CAGAACCTCG CCCAGCGCCC 60
      ACCATGCCCC GGCAGCTCAG CCGCGCGGCC GCGCTCTTCG CGTCCCTGGC CGTAATTTTG 120
      CACGATGGCA GTCAAAATGAG AGCAAAAGCA TTTCCAGAAA CCAGAGATTA TTCTCAACCT 180
65     ACTGCAGCAG CAAACAGTACA GACATAAAA AAACCTGTCC AGCAACGAGC TAAGCAAGCA 240
      CCTCACCAAA CTTTAGCAGC AAGATTCAAT GATGGTCATA TCACCTTTCA AACAGCGGCC 300
      ACAGTAARAA TTCCAACAAC TACCCAGCA ACTACAAAA ACCTGCAAC CACCGGCCA 360
      ATTACTTACA CCTGTGTCAC AACCCAGGCC ACACCAACA ACTCACACAC AGCTCTCTCA 420
      GTTACTGAAG TTACAGTCCG CCTAGCTTA GCGCTTATT CACTGCCACC CACCATCACC 480
70     CCACGAGCTC ATACAGCTGG AACCAATTCA TCAACCGTCA GCCACACAAC TGGGAACACC 540
      ACTCAACCCA GTAACAGAC CACCTTCCA GCAACTTTAT CGATAGCACT GCACAAAGC 600
      ACAACCGTTC AGAAGCCTGA TCACCCACC CATGCCCCAG GAACAACGCG AGCTGCCAC 660
      AATACACCCC GACACCTGCG AACTGCTCC ACGGTTCTCG GCGCCACCTT TGCACCTCAG 720
      CCATGTCAG TCAAGACTGG AATTATCAG GTTCTAAACG GAAGCAGACT CTGTATAAAA 780
75     GCAGAGTAGG GGTATACAGT GATTGTTCAA GACAAGGAGT CGGTTTPTTC ACCTCGGAGA 840
      TACTTCAACA TCGACCCCAA CGCAACGCAA GCTCTCGGGA ACTGTGCAC CCGAAATCC 900
      AACCTCTGT TGAATTTTCA GGGCGGATTT GTGAATCTCA CATTTACCA GGATGAAGAA 960
      TCATATTATA TCAATGAAGT GGGAGCTTAT TTAGCGTCTT CAGATCCAGA GACAGTTTAC 1020
      CAGGAATCA AACATGCGGT GGTGATGTTT CAGACAGCAG TCGGCGATTC CTTCAGTGC 1080
80     GTAGTGAAC AGAGCTTCCA GTTGTACGCC CACCTGAGG TGAACAACAC CGATGTCCAA 1140
      CTTCAAGCTT TTGATTTTGA AGATGACCAC TTTGGAAATG TGGATGAGTG CTGCTCTGAC 1200
      TACACAATTG TGCTTCTCTG GATTGGGGCC ATCGTGGTTG GTCCTGCTT TATGGGTATG 1260
      GGTGTCTATA AAATCCGCTT AAGGTGTCAA TCATCTGGAT ACCAGAGAA CTAAATGTTG 1320
      CCCGGGGGGA ATGAAAATAA TGAATTTAG AGAATCTTT CATCCCTTCC AGGATGGATG 1380

```

5
 10
 15
 20
 25
 30

```

TTGGGAAATT CCCTCAGAGT GTGGGTCCTT CAAACAATGT AACCACCAT CTTCATTCA 1440
AATGAAGTGA GTCATGTGTG ATTTAAGTTC AGGCAGCACA TCAATTTCTA AATACCTTTT 1500
GTTTATTTTA TGAAGATAT AGTGAGCTGT TTTTCTTTA GTTTCCTTTA GAATATTTTA 1560
GCCACTCAA GTCAACATTT GAGATATGTT GAATTAACAT AATATATGTA AAGTAGAATA 1620
AGCCTTCAA TTATAAACCA AGGGTCAATT GTAACFAATA CTACTGTGTG TGCATTGAAG 1680
ATTTTATTTT ACCCTTGATC TTAACAAGC CTTTGCTTTG TTATCAAATG GACTTTCAAT 1740
GCTTTTACTA TCTGTGTTT ATGGTTTCAT GTAACATACA TATTCCTGGT GTAGCACTTA 1800
ACTCCTTTTC CACTTTAAAT TTGTTTTTGT TTTTGTAGAC GAGGTTTCAC TCTGTCCACC 1860
CAGGCTGGAG TACAGTGGCA CGATCTCGGC TTATGGCAAC CTCGCGCTCC CGGGTTCAAG 1920
TGATTCTCCT GCTTCGCTT CCGAGTAGC TGGGATTACA GGCACACACT ACCACGCTG 1980
GCTAATTTTT GTATTTTTAT TATAGACGGG TTTCACCATG TTGGCCAGAC TGGTCTTGAA 2040
CTCTTGACCT CAGGTGATCC ACCCACCCTA GCCTCCCAA GTGCTGGGAT TACAGGCATG 2100
AGCCATTGGC CCGGCGCTTA AATGTTTTT TTAATCATCA AAAAGAACA CATATCTCAG 2160
GTGTCTAAG TGTTTTTATG TAAACCAAC AAAAGAACA AATCAGCTTA TATTTTTTAT 2220
CTTGATGACT CCGCTCCAG AATGTCTAGA CTAAGAATTA GGTGGCTACA GATGGTAGAA 2280
CTAAACAATA AGCAAGAGAC AATAAATATG GCCCCTAAT ATTACAAG TGCCAGAGTC 2340
TAGGCTAAGC ACTTTATCTA TATCTCATTT CATTCCTACA ACTTATAAGT GAATGATTA 2400
ACGTAGACTT AAGGGAACCT AATCACTTAA ATGTCACTG GCTAAGCTAT GGCAGAGCCA 2460
GAGCTTGAAT TCAGTTGGT CTGACATCAA GGTCTTTGGT CTTCTCCCTA CACCAAGTTA 2520
CCTACAAGAA CAACTGACCC ACCTCTGCC TGAAGGCTCA CACCTCATAC CAGCATACGC 2580
TCACCTTACA GGGAAATGGG TTTATCCAGG ATCATGAGAC ATTAGGGTAG ATGAAAGGAG 2640
AGCTTTGCGG ATACCAAAAT AGCCTATCCT TAATAAATCC TCCACTCTCT GGAAGGAGAC 2700
TGAGGGGCTT TGTAAAACAT TAGTCAGTTG CTCATTTTAA TGGGATTTCT TAGCTGGGCT 2760
GTAAAGATGA AGGCATCAA TAAACTCAA GTATTTTTAA ATTTTTTTGA TAATAGAGAA 2820
ACTTCCTCAA CCACTGCTT TTTCTGAGT GTATAGCCCT ATCTGTGGT AACCTGTCTG 2880
TTCTGCACCT CATATCCATA TTTCTATTG TTCACCTTAT TCTGTAGAGC AGCCTGCCAA 2940
GAATTTTTAT TCTGTGTTT TTTTGTCTG TAAAGRAAG AACTAAGTCA GGATGTTAAC 3000
AGAAAAGTCC ACATAACCT AGAATCTTA GTCAAGGAAT AATTCAGTCA AGCCTAGAGA 3060
CCATGTTGAC TTCTCTCATG TGTTCCTTA TGACTCAGTA AGTTGGCAAG GTCTTGACTT 3120
TAGTCCTAAT AAAACATTGA ATTGTAGTAA AGTTTTTTC AATAAAAACT TACTTTGG
  
```

Seq ID NO: 230 Protein sequence

Protein Accession #: NP_055213.1

35
 40
 45

```

1 11 21 31 41 51
| | | | |
MPRLSAAAA LFASLAVILH DGSQMRKAF PETRDYSQPT AAATVQDIKK PVQPAKQAP 60
HQTLAARFMD GHITPQTAAT VKIPTITPAT TKNTATTSPY TYTLVTQAT PNNSHAPPV 120
TEVTVGSLA FVSLPPTIIP PAHTAGTSSS TVSHTTGNTI QPSNQTLPA TLSLALKST 180
TGQKPDQPTH APGTTAAAEH TERTAAPAST VPGPTLAPQ SSVKGTIYQV LNSRLCIKA 240
EMGIQLIVQD KESVSPRRY FNIIDPNATQA SENCGRKSN LLLNFGGSFV NLTFTKDEES 300
YYISEVGAYL TVSDPSTVYQ GIKRAVVMFQ TAVGHSFKCV SEQSLQLSAH LQVKYTDVQL 360
QAFDFEDDFH GNVDCSSDY TIVLPVIGAI VVGLCLMGNG VYKIRLRCSG SGYQRI
  
```

Seq ID NO: 231 DNA sequence

Nucleic Acid Accession #: NM_005409.3

Coding sequence: 94..378

50
 55
 60
 65
 70
 75

```

1 11 21 31 41 51
| | | | |
TTCCTTTTAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAGGCT GAGTAGCAG 60
CAACAGCACCC AGCAGCAACA GCAAAAAACA AACATGAGTG TGAAGGCGAT GGCTATAGCC 120
TTGGCTGTGA TATGTGTGTC TACAGTTGTT CAAGGCTTCC CCATGTTCAA AAGAGGAGCC 180
TGTCTTTGCA TAGGCCCCGG GGTAAAGACA GTGAAAGTGG CAGATATTGA GAAAGCCTCC 240
ATAATGTACC CAGTAAACAA CTGTGACAAA ATAGAAGTGA TTATTACCTT GAAAGAAAT 300
AAGAGACAAC GATUCCTAAA TCCCAATCG AAGCAAGCAA GGCTTATAT CAAAAAGTT 360
GAAGGAAGA ATTTTTAAAA ATATCAAAAC ATATGAAGTC CTGGAAGAG GCATCTGAAA 420
AACCTAGAAC AAGTTTAACT GTGACTACTG AAATGACRAG AATCTACAG TAGGAAACTG 480
AGACTTTTCT ATGGTTTGT GACTTTCAC TTTGTACAG TTATGTGAG GATGAAAGGT 540
GGGTGAAAGG ACCAAAAACA GAAATACAGT CTTCTGAAAT GAATGACAA CAGAATTCCA 600
CTGCCCAAAG GAGTCCAGCA ATTAATGGGA TTTCTAGGAA AAGCTACCTT AAGAAAGGCT 660
GGTTACCATC GGAGTTTACA AAGTGCTTTC ACGTTCTTAC TTGTTGTATT ATACATTCA 720
GCATTTCTAG GCTAGAGAAC CTTCTAGATT TGATGCTTAC AACTATTCTG TTGTGACTAT 780
GAGAACATTT CTGTCTCTAG AAGTTATCTG TCTGTATTGA TCTTTATGCT ATATTACTAT 840
CTGTGTTTAC AGTGGAGACA TTGACATTAT TACTGAGTTC AAGCCCTTAT AAGTCAAAAG 900
CATCTATGTG TCGTAAGACA TTCTCAAAAC ATTTTTCAT CCAATACAC ACTTCTTTCC 960
CCAAATATCA TGTAGCACAT CAATATGTAG GGAACATTC TTATGCATCA TTTGGTTTGT 1020
TTTATAACCA ATTCATTAAA TGTAAATCAT AAAATGTACT ATGAAAAAAA TTATAAGCTA 1080
TGGGATACAG GCAACAGTGC ACATATTCCA TAACCAAAAT AGCAGCACCG GTCTTAATTT 1140
GATGTTTTTC AACTTTATT CATTGAGATG TTTTGAAGCA ATTAGGATAT GTGTGTTTAC 1200
TGTACTTTT GTTTGTATCC GTTTGTATAA ATGATAGCAA TATCTTGGAC ACATTGAAA 1260
TACAAATATG TTTGTCTAC CAAGAAAAA TGTGAAAAA TAAGCAATG TATACCTAGC 1320
AATCACTTTT ACTTTTTGTA ATTCGTCTC TTAGAAAAAT ACATAATCTA ATCAATTTCT 1380
TTGTTTATGC CTATATACCTG TAAATTTAG GTATACTCAA GACTAGTTTA AAGAATCAA 1440
GTCATTTTTT TCTCTAATAA ACTACCACAA CCITTCTTTT TAAAAAAA AAA
  
```

Seq ID NO: 232 Protein sequence

Protein Accession #: NP_005400.1

80

```

1 11 21 31 41 51
| | | | |
MSVKGMAIAL AVILCATVVQ GPFMPKRGRC LCIGPGVKAV KVADIEKASI NYPNNCDKI 60
EVIITLKENK GQRCINPKSK QARLIKKIVE RKNF
  
```

Seq ID NO: 233 DNA sequence
Nucleic Acid Accession #: NM_000577.1
Coding sequence: 41..520

```

5      1      11      21      31      41      51
|      |      |      |      |      |
GGCAGGAGGG GAAGACCTCC TGTCTATCA GGCCTCCCC ATGGCTTAG AGACGATCTG 60
CGAGCCCTCT GGGAGAAAT CCAGCAAGAT GCAAGCCTTC AGAATCTGG ATGTTAAACA 120
GAAGACCTTC TATCTGAGGA ACACCAACT AGTTGCCGA TACTTGCAAG GACCAAATGT 180
10 CAATTAGAA GAAAAGATAG ATGTGGTACC CATTGAGCCT CATGCTCTGT TCTTGGGAAT 240
CCATGGAGGG AAGATGTGCC TGTCTGTGT CAGTCTGGT GATGAGACCA GACTCCAGCT 300
GGAGGCAGTT AACATCACTG ACCTGAGCGA GAACAGAAAG CAGGACAAGC GCTTCGCCTT 360
CATCCGCTCA GACAGTGGCC CCACCACCAG TTTTGAGTCT GCGGCTGDC CCGGTTGGTT 420
CCTCTGCACA GCGATGGAGG CTGACCCGCC CGTCAGCCTC ACCAATATGC CTGACGAAGG 480
CGTCATGGTC ACCAAATTCT ACTTCCAGGA GGAAGAGTAG TACTGCCCAg GCGTGCCTGT 540
15 TCCTATTCTT GCATGGCAAG GACTGCAGGG ACTGCCAGTC CCCCTGCCCC AGGGCTCCCG 600
GCTATGGGGG CACTGAGGAC CAGCCATTGA GGGGTGGACC CTCAGAAGGC GTCACAACA 660
CCTGGTCACA GGACTCTGCC TCCTCTTCAA CTGACCCAGC TCCATGCTGC CTCAGAATG 720
GTCTTTCTAA TGTGTGAATC AGAGCAGAGC AGCCCTGCA CAAAGCCCTT CCATGTGCGC 780
20 TCTGATTAAG GATCAAAACC CCGACCCCTT GCGCAACCTG CTCTCTCTTT GCGACTGCGT 840
CTTCTCCCTT CATTCACCTT TCCCATGCC TGGATCCATC AGGCCACTTG ATGACCCCCA 900
ACCAAGTGGC TCCCACACCC TGTPTTACAA AAAAGAAAAG ACCAGTCCAT GAGGGAGGTT 960
TTTAAGGGTT TGTGGAAAAT GAAAATTAGG ATTCATGAT TTTTATTTT CAGTCCCCGT 1020
GAAGGAGAGC CCTTCATTTC GAGATTATGT TCTTTCGGGG AGAGGCTGAG GACTTAAAT 1080
25 ATTCTGCTG TGTGAATG ATGTGAAAG TAAAGTGGT CTTTCCCTT CTTTTCCTC 1140
TTTTTTTGG ATGTCCCAAC TTGTAAAAAT TAAAGTTAT GGTACTATGT TAGCCCATGA 1200
ATTTTTTTT TCCTTTTAAA ACATCTCCAT AATCTGGACT CCTCTGTCCA GGCAGTCTG 1260
CCAGCCCTCC AAGCTCCATC TCCACTCCAG ATTTTTTACA GCTGCTGCA GTACTTTACC 1320
TCCTATCAGA AGTTTCTCAG CTCACAGGCG TCTGAGCAAA TGTGGCTCCT GGGGGTTCTT 1380
30 TCTTCTCTG CTGAAGGAAT AAATTCCTCC TTGACATTGT AGAGCTTCTG GCATCTGAG 1440
ACTGTATGA AAGATGGCTG TGCTCTGCC TGCTCCCCC ACCAGGCTGG GAGCTCTGCA 1500
GAGCAGAAAC CATGACTCGT ATATGTCTCA GGTCCCTGCA GGCACAGCA CCTAGCTCG 1560
CTCTTGGCAG GTACTCAGCG AATGAATGCT GTATATGTTG GGTGCAAGT TCCCTACTTC 1620
35 CTGTGACTTC AGCTCTGTTT TACATAAAA TCTTGAAAT GCCTAAAAA AAAAAAAA 1680
AAAAAAA AAAA AAAA AAAA

```

Seq ID NO: 234 Protein sequence
Protein Accession #: NP_000568.1

```

40      1      11      21      31      41      51
|      |      |      |      |      |
MALETCRPS GRKSSKMQAF RIWDVNQKTF YLRNQLVAG YLQGENVME EKIDVVPFEP 60
EALFLGHHG KMCLSCVKSQ DETRLGLEAV NITDLSENK QDKRPAFIRS DSGPTTSFES 120
45 AACPGWFLCT AMEADQPVSL TNMPDEGVMV TKFYFQED

```

Seq ID NO: 235 DNA sequence
Nucleic Acid Accession #: NM_001840.1
Coding sequence: 149..1567

```

50      1      11      21      31      41      51
|      |      |      |      |      |
GGGGACTACG GAGAGCTCTG CAGGGAGCGG AGGCCCCCGC CCGGGCCAAG GGAGCTCTG 60
TCCGAGGAGC CAGGGGATGC GAAGGGAATG CCCCTGTGG GTCACTTCTT CAGTCATTTT 120
GAGCTCAGCC TAATCAAGA CTGAGGTTAT GAAGTCGATC CTAGATGGCC TTGCAGATAC 180
55 CACCTTCGCG ACCATCAACA CTGACCTCCT GTACGTGGGC TCAAATGACA TTCAGTACGA 240
AGACATCAAA GGTGACATGG CATCCAAAT AGGGTACTTC CCACAGAAAT TCCCTTTAAC 300
TTCTTTTAGG GGAAGTCCCT TCCAGAGAA GATGACTGCG GGAGACAACC CCCAGCTAGT 360
CCACGAGAC CAGGTGAACA TTACAGAAT TTACAACAG TCTCTCTCTT CCTTCAAGGA 420
GAATGAGGAG AACATCCAGT GTGGGGAGAA CTTCATGGAC ATAGAGTGTG TCATGGTCTT 480
60 GAACCCGAGC CAGCAGCTGG CCATTGCACT CTTGTCTCTC ACCTGAGGCA CCTTCACGGT 540
CCAGGAGAAC CTCTGTGTGC TGTGGTTCAT CCACTACTCC CGCAGCTCC GCTGCAGGCC 600
TTCCTACCACT TTCTATGGCA GCCTGGCGGT GGCAGACCTC CTGGGAGTG TCATTTTGT 660
CTACAGCTTC ATTGACTTCC ACCTGTTCCT CCGCAAGAT AGCCGCAACG TGTTCCTGTT 720
CAAATCGGGT GGGGTACGGG CCTCTTCAC TGCCCTCCGTG GGCAGCTGTG TCCTCAGAGC 780
65 CATCGACAGG TACATATCCA TTCACAGGCC CCTGGCCTAT AAGAGGATTG TCACAGGCC 840
CAAGGCGGTG GTGGCGTTTT GCCTGATGTG GACCATAGCC ATTGTGATG CCGTGTCTGC 900
TCTCTGTGGC TGAACCTGCG AGAACTGCA ATCTGTTTGC TCAGACATT TCCACACAT 960
TGATGAAACC TACCTGATGT TCTGGATCGG GGTCAACAGC GTACTGCTTC TGTTCATCGT 1020
GTATGCTTAC ATGTATATTC TCTGGAAGGC TCACAGCCAC GCGTTCGCA TGATTCAGCG 1080
70 TGGCAACCAG AAGAGCATCA TCATCCACAC GTCTGAGGAT GGGAAAGTAC AGGTGACCCG 1140
CGCAGACCAA GCCCGCATGG ACATTAGGTT AGCCAAAGAC CTGGTCTCTG TCCTGGTGGT 1200
GTTGATCATC TGTGGGGGCC CTCTGCTTGC AATCATGGTG TATGATGTCT TTGGGAAGAT 1260
GAACAAGCTC ATTAAGACGG TGTGTGCTAT CTGCAGTATG CTCTGCTGCG TGAACCTCAC 1320
CGTGAACCCC ATCATCTATG CTCTGAGGAG TAAGGACCTG CGACACGCTT TCGGAGCAT 1380
75 GTTTCCTCT TGTGAAGGCA CTGCGAGGCC CTGATATAAC AGCATGGGGG ACTGGAGCTG 1440
CCTGCACAAA CACCAACACA ATGCAGCCAG TGTTCACAGG GCGCAGAAAC GCTGCATCAA 1500
GAGCAGGGTC AAGATTGCCA AGGTAAACAT GTCTGTGTCC ACAGACAGT CTGGCAGGCC 1560
TCTGTGAGCC TGATGCCCTC CTGGCAGCAC AGGAAAGAA TTTTATTTT TAAGCTCAA 1620
80 ATCTAGAGGA GTCTATTGTC TCCTTGTGTA TATTTTATTA ACTTTACCAT GCTCAATGAA 1680
AAGGTGATTG CCATGATGCA CTATTGTGCT TAGTTTCCGT TTGGGCTPAT CTTCCGGGGT 1740
TGATAGGAAA CCTTT

```

Seq ID NO: 236 Protein sequence
Protein Accession #: NP_001831.1

	1	11	21	31	41	51	
5	MKSILDLGLAD	TFRTITITDL	LYVGSNDIQY	EDIKGDMAK	LGYPFQKFPL	TSFRGSPFQE	60
	KMTAEDNPQL	VPADQVNITE	FYNKSLSEFK	ENEENIQCGE	NFMDIECPMV	LNPSQQLALA	120
	VLSLTLGTPT	VLENLLVLVC	ILHSRLSLRCR	PSYHFIGSLA	VADLLGSVIF	VYSFIDPHVF	180
	HRKDSRNVPFL	FKLGGVTASF	TASVGSFLFT	AIDRYISIH	PLAYKRIVTR	PKAVVAFCLM	240
	WTIAIVIAVL	PLLGWNCERL	QSVCSDIFFH	IDETYLMEFI	GVTSLVLLPI	VYAYMYILWK	300
	AHSHAVRMIO	RGTQKSIH	TSRDGKVQVT	RPDQARMDIR	LAKTLVLILV	VLITCWGPLL	360
10	AIMVYDVFGK	MNKLKTVFA	FCSMLCLLNS	TVNPIIYALR	SKDLRHAFRS	MEFSCBETAQ	420
	PLDNSMGDSO	CLKHANNAA	SVHRAAESCI	KSTVKIAKVT	MSVSTETSAE	AL	

Seq ID NO: 237 DNA sequence

Nucleic Acid Accession #: NM_016083.2

Coding sequence: 64..1482

	1	11	21	31	41	51	
20	GATTGCCCCC	TGTGGGTAC	TTTCTCAGTC	ATTTTGAGCT	CAGCCTAATC	AAAGACTGAG	60
	GTTATGAAGT	CGATCCTAGA	TGGCCITGCA	GATACCACTT	TCCGACCACT	CACCACTGAC	120
	CTCCCTGTACG	TGGGCTCAAA	TGACATTGAG	TACGAAGACA	TCAAAGGTGA	CATGGCATCC	180
	AAATTAGGGT	ACTTCCACAC	GAAATTCCTT	TAACTTCTCT	TAGGGGGAAG	TCCCTTCCAA	240
	GAGAAGATGA	CTGCGGGAGA	CAACCCCCAG	CTAGTCCGAG	CAGACCAGGT	GAACATTACA	300
	GAAATTTACA	ACAAGTCTCT	CTCGTCCCTT	AAGGAGAAATG	AGGAGAACAT	CCAGTGTGGG	360
25	GAGAACTTCA	TGACATAGAA	GTGTTTCATG	GTCTTGAACC	CCAGCCAGCA	GCTGGCCATT	420
	GCAGTCCCTG	CCCTCAGGCT	GGGCACCTTC	ACGGTCCCTG	AGAACCTTCT	GGTGTCTGTG	480
	GTCTATCTCC	ACTCCGCGAG	CTCCGCTGCG	AGGCTTCTCT	ACCACTTCAT	CGGCGAGCTG	540
	GCGGTGGCAG	ACCTCCTGGG	GAGTGTCAAT	TTTGTCTACA	GCCTTCATTGA	CTTCCACGTG	600
	TTCCACCGCA	AAGATAGCCG	CAACGTGTTT	CTGTTCAAA	TGGGTGGGGT	CACGGCTTCC	660
30	CTCACTGCCCT	CCGTGGCGAG	CTGTTCCTTC	ACAGCCATCG	ACAGGTACGT	ATCCATTGAC	720
	AGGCCCTTGG	CCATATAGAG	GATGTGACCC	AGGCCCAAGG	CCGTGTGTGG	GTTTTGCTCG	780
	ATGTGGACCA	AGCCCATGCG	GATGCGCGTG	CTGCTCTCTC	TGGGCTGGAA	CTGCGAGAAA	840
	CTGCAATCTG	TTTGCTCAGA	CATTTTCCCA	CACATTGATG	AAACCTACCT	GATGTCTCGG	900
	ATCGGGGTCA	CCAGCTTACT	GCTTCTGTTC	ATCGGTGATG	CGTACATGTA	TATTCTCTGG	960
35	AAGGCTCACA	GCGATGCGAT	CCGCAATGAT	CAGCGTGGCA	CCAGAGAGAG	CATCATCATC	1020
	CACAGCTCTG	AGGATGGGAA	GGTACAGGTG	ACCGCGCCAG	ACCAAGCCCG	CATGGACATT	1080
	AGGTAGAGCA	AGACCTTGGT	CTGATCTCTG	GTGGTGTGTA	TCATCTGCTG	GGGCCCTTTG	1140
	CTTGCAATCA	TGGTGTATGA	TGCTTTTGGG	AAGATGAACA	AGCTCATTTA	GACGGTGTCT	1200
	GCATTCGTCA	GTATGCTCTG	CCTGTCTGAAC	TCCACCGTGA	ACCCCATCAT	CTATGCTCTT	1260
40	AGGATAAGG	ACCTGGGACA	CGCTTTCGGG	AGCATGTGTC	CCTCTGTGTA	AGGCACTCGG	1320
	CAGCCTCTGG	ATAACAGCAT	GGGGGACTCG	GACTGCTCTG	ACAAACAAGC	AAACATGCA	1380
	GCCATGTGTC	ACAGGGCGCG	AGAAAGCTGC	ATCAGAGACA	CGGTCAAGAT	TGCCAAGGTA	1440
	ACCATGTCTG	TGTCCACAGA	CAGTCTGCGC	GAGGCTCTGT	GAGCCTGATG	CCTCCCTGGC	1500
45	AGCACAGGAA	AAGATATTTT	TTTTTTAAGC	TCAAAATCTA	GAGAGTCTA	TTGTCTCTCT	1560
	GTTATATATT	TTTTAACTTT	ACCATGCTCA	ATGAAAAGGT	GATGTGCACT	ATGATCACTT	1620
	ATCAGTTTGC	TAATGTCTCC	ATAGTITTAG	TACTCAAACT	CCATTCTCCA	GGGGTTTACA	1680
	GTGAAGAAAG	CTGTTTGTCT	AAGTGACTGA	ACGATCCTTC	AAAGCTCAA	TGAATAGGTA	1740
	GGGAAACCTT	TGGCTACACA	ATTGGAAGTC	TAAGAACCCA	TGGAAAATG	CCATCAAATG	1800
50	ATATATGCTT	TTGTAACCC	AACTTTTCACT	ATAATGTGAA	ATGTAACCTG	CCGTATGATC	1860
	AGAGATGTCC	ATTTTATCAA	GTTATAGTAC	TAGAGATATT	TGTAAAATG	TATTATGTCC	1920
	TGTAGATGTG	GTATCAGTGT	TTATGTGCTA	TTAATATTGG	TTTATGTTAG	CAAAACAGAA	1980
	AGGTAGACTT	TTATGAGATC	AATGGACAAG	CAGTGGATAC	GTGTCAATGT	GTGCACTTTT	2040
	TTTCTATATT	ATTGCCCATG	ATATTAATTT	AGAAATAAAC	CTTAATATTT	CTTCAATAT	2100
	CTCTATTTAA	TTTTCAGACT	GAAATAACCG	TAAAGGTTTA	TTTTTCTGTT	ACCTCAACAA	2160
55	GAAGAATTGG	AAGACTTCAA	AATATTGAGC	AGAATTCATT	CATACCTAAA	AATTTATTAG	2220
	CCCTGCATT	TCAATAGGAG	ACACATTATC	TTCTGGACTA	TAGCTGTCTT	ATATGGATTAT	2280
	AATCAGATG	GAGAGAGGAA	AGCATATTGA	CTTTTTTTGA	GCGACATCTC	TGACTTTCTT	2340
	TAGTCTTTAG	CTATTACTGG	ATCTCTTAAG	ACAGCATGTG	TTAATCTTAA	TGTATATCGT	2400
60	TATCACTGTG	CAGTGTCTGT	TTACTTGAAT	AGTATTGTGT	TOCTATATTC	CAGGTTTAG	2460
	TAGATTTCACT	GCCTGGGTGG	CCAAACAACA	GTCTTCATT	TTTTTAATG	AAAAGAAATA	2520
	GTGTCTGGAT	CAGTAAATTT	ATACTGTGTG	TGAGTGTGAA	TATAAATGTG	TGTATGTGTG	2580
	TTTCTGTCCG	TAACTGTTC	AGTAATGTCA	TAAAGTGAGA	AACTGTGAC	CAGTATAAAA	2640
	CTTTTACCAC	TGCTGTCACT	CTTCCACATG	GATTCAGTTT	CIAAAATGGA	GTCTCTCTCT	2700
65	TAATCTTGTT	GATAAAANTA	CTGACTCCAA	CCATTCAAAA	ATTTCAACCC	ATCCCTCTCT	2760
	AAGAGATTGG	ATCAAGTATT	ACTAAATTGA	CCTTTAGGTA	TTACACAAGA	CCAGTCTTAA	2820
	GCAAAAATA	ATGACAGGCA	TCCAAAGGAG	GGATGTATTT	TGATGTGTAT	TGCCAGGAAA	2880
	GGAGAGTACT	TTGGTTTCTG	AGCAACGAAT	ATTGAGCAAT	ATGTCAGTCA	CTAAAAGGAA	2940
	GACAGTCTTA	CAGAAAAACA	AATGGTAACA	TTTTTCAATA	GCGTGTGTAG	ATAGTATGCA	3000
70	CTATATACAT	CACGTTAAG	TAGSACTATC	ACACCCAGCC	CATGTGGCTA	AAAAAGCTGA	3060
	ATCAGACAGT	GGATGAGACA	CACAAAGGCA	GTGAAGAAC	GATACACTTG	GCATTGACGT	3120
	CTAGCTATGC	TGTATCTGTG	CTTTGCCAC	ATGCCCTTGG	TGACAGCTGA	GCACCCAGCT	3180
	CTGTCTTGGT	AGCTTGTGGC	TAAAGAACAA	ATCTCTCTCT	TGCTGTGTGT	TAGCAAGATA	3240
	CACCTAAGCA	TGAAGATAAA	CACAGCTGCT	TTCTTCTTAC	ACCCCGGTCT	CATGCTCTCT	3300
75	AATGGGCGCA	TGGTGTCTTG	TTGGGCTTTT	TTCCAGTAAG	GAATGATATT	GCTGAAGAA	3360
	CTACTTAACC	CTGACAAATT	TTAATTATTA	TTCTTCTTAA	TACAGATAAA	ACATGCTTCC	3420
	TACAAAGCCC	CAAGGTTTAC	ATAGTCTGAA	GTGAAGTACA	GAGCTGGCAT	CTATCTGTGT	3480
	ATTCTAGATC	CTCAGATAC	CCAAGCAGCC	TGAAGGGGCA	GTCTCTCTCT	TTACGGTTCA	3540
	CGCTCTAAGG	CAGGATGTGG	CTTATGAGAT	ACITTGCAAT	GTCTGTCTGC	ACACCTTGAA	3600
80	TCTGCCCTGT	GGCTCCCTTA	CTTACCTCT	CTGTCTATGT	CAGATGAAGG	CTCAGGGTGC	3660
	TAGAGGATTA	CTAAGATCTC	TTTCTAAAGA	CAGGAGAGAT	TATTTACAAG	AAGAACTCAC	3720
	CAGGGTTTAG	TTTGCAATTA	AGAATTGCCA	GTCTTTTGTG	CTGCATCATC	TTGAACATTA	3780
	ATCCACATGT	TTACAGCTC	ACCAGGCAGT	ACCAATGCTC	TTTTACAGC	TATGAAGAGC	3840
	TAGAGAAATT	CTGTATATGG	TAGAAAAATT	TCACGGTTCA	TTTTTGAAGC	TGCATTTGTG	3900
	CGTATGCAGT	GTAGATTTTA	TAGTGTGTG	TGCTTTCAAG	AICPAATCA	TATATATATA	3960

5 ATTAAGGGAC AATGGGGCTG ACAGCACTAA ACTTGGTGCT TATTGATATT CTAAGAAATA 4020
 TCTGGGAAAT ATCATACGTT ATGTTATACA ACCTTCATTT AAAAAGGTTT AAAACTAGTT 4080
 AGATTCACTT TGACACTTTT CATATCATTI CTTAACCCAA GTGACGAAAA CATTGTCCCC 4140
 AATGAATATA CTCATTAGAA TTACCATTTG TTAATATCAC TCATTAATTA ACCCCATAAT 4200
 TAGATCCATT AATTAAATG ATTABATTT AAGTAGTTT TATAAGGTCT GACATCAGAG 4260
 GTATCTTACT TTCTCTGAG GATGATGTAC TTGCCCTGAC CATGCATTTT ACCATCACAC 4320
 ATGTTTCAAG AAGGCCAAAT TCCCAACCTG CTCATITTTT TTTTATCAG AGTCATGATG 4380
 AATCAGTCCT AGAATGTTT ATTTCACAA GTAGGGCTGC CTCCAAGAGG AACCTCTGAT 4440
 TTATTTTGTA TGAATATAT GTGAAGGAT ATGAATCTGA GAGATGCTGT AGACATCTGT 4500
 10 CCTACACTTG AGATGATTTT CAAGCCTCTC TGGCACTTTG AGTTAAGTCT ATCTGGTATT 4560
 AAATGCCAAG GACCTTTTGC TGCCATAATC CACTCTGCGG GAAATAGGCC CAACCCACAG 4620
 ATGAGAATTA GGCCCTGGAT GAGTAGCGCT ATAGTTACTG TCCGTGTGAT TAATTTCTGC 4680
 CATTTCATGT CCATAAAGA GACCACCCAT ATCATGCACA CAATTAGATT TCTCACACTC 4740
 TAACGTATA TTTGTATGAT ATTTTAAAT CTCTTAAATG CTGGGCAATG GCTATTAAAC 4800
 15 ATTAATTGTC TTGCACTGEC CTTCGTATGA AATGTAAAC ATGCTTATTG TAATATAGAA 4860
 AAAAACATTG TATCTACTGA TTTGGGCTGA ATGTATGTAA ATAGGTTTCT AAAAAGTCAG 4920
 ATGTTTGAGC AGTGCCCTAC AAATCAGTAA TTTTCGGGTG GGAGAGTTTC TTTACATTGC 4980
 CGTGGCATCT TAAAAGCTTAT CTTCATGTAA ATTGACTGTA CTAGGCCCTAC TGGGGATCAG 5040
 AGTTCCCAAG AAAGGAAACC TTTTCTGTGA TCTGGATTCA AATTATTATC CAATGTTTCA 5100
 20 AGCGGGAAAC ATGACTCTTT ATTGTCTGTA AATCTAACAT TATTACTTTT CCTCTTAGAA 5160
 GAATATTGTA TTGTAGATG TTTGTGAGC TGGTAACATC GTTGCAACCA CTGCAATATC 5220
 TTGTTAGTA ATCTGTATAA TACTTTGTAT ACAAGTACTG GTAAGATTGT TATTAAATGT 5280
 AGCTTCAGTC ATTAATATAC TATAGCAAAG TAGTACTTCT TCTGTAATAT TTACAATGTA 5340
 TTAAGCCAC AGTATATTTT ATTTCAATGT AATTAACATG TTAACTTATT CAAAGAGAAA 5400
 25 ACATCTCATC ATGCTCTATTG TCCAAAGTGA CTTGGAAATCA AATAAAATTT CTAGATTACC 5460
 ATGAAGAACA TA

Seq ID NO: 238 Protein sequence
Protein Accession #: NP_057167.1

30 1 11 21 31 41 51
 MKSILDGLAD TFRITITDL LYVGSNDIQY EDIKGDMASK LQYFPQKFFL TSFRGSPFQE 60
 35 KMTAGDNEQL VPADQVNITE FYNKSLSSEK ENEENIQCGE NFMDIECFMV LNPSQQLAIA 120
 VLSLTGLTFT VLENLVLVLCV ILHRSRLRCR PSYHFIGSLA VADLLGSVIF VYSPIDFHV 180
 HRKDSRNVL FKLGGVTASL TASVGSLEFIT AIDRYVSIHR PLAYKRIVTR KAVVAFCLM 240
 WTIALAVL PLEGNWCEKL QVCSNDFPH IDETYLMFWI GVTSLVLLFI VYAYMYLWK 300
 AHSBAVMRIQ RGTQKIIH TSDEGKVQVT RPDQARMDIR LAKTLVLLIV VLIICWGPLL 360
 40 AIMVYDFGK MNKLKIVFA FCSMLCLNS TVNPIIYALR SKDLRHAFRS MFPSCBETAQ 420
 PLDSMGSDD CLKHANMAA SVHRAESCI KSTVKIAKVT MSVSTDTSAE AL

Seq ID NO: 239 DNA sequence
Nucleic Acid Accession #: NM_033181.1
Coding sequence: 17..1252

45 1 11 21 31 41 51
 ATGAAGTCGA TCCFAGATGG CCTGCGAGAT ACCACCTTCC GCACCATCAC CACTGACCTC 60
 50 CTGTACGTTG GCTCAAATGA CATTCACTAC GAAGACATCA AAGGAGAATG AGGAGAACAT 120
 CCAGTGTGGG GAGAACTTCA TGGACATAGA GTGTTTCATG GTCCGGAACC CCAGCCAGCA 180
 GCTGGCCATT GCAGTCTCTT CCTCACGCT GGGCAOCTTC ACCTGCTCTG AGAACCTCTT 240
 GGTGCTGTGC GTCATCTTCC ACTCCGCGAG CCTCCGCTGC AGGCCTCTCT ACCACTTCAT 300
 CGGCGAGCTG CGCGTGGCAG ACCCTCCGGG GAGTGTCAAT TTTGTCTACA GCTTCATTGA 360
 55 CTCCACGCTG TTCCACGCGA AAGATAGCGG CAACGTGTTT CTGTTCAAAC TGGGTGGGGT 420
 CACGCGCTCC TTCACTGCTT CCGTGGCGAG CCGTTCCTC ACAGCCATCG ACAGGTACAT 480
 ATCCATTGAC AGGCCCCCTGG OCTATAAGAG GATTGTCAAC AGGCCCAAGG CCGTGTGTGC 540
 GTTTTGCTTG ATGTGGACCA TAGCCATTGT GATGCGCGTG CTGCTCTCC TGGGCTGGAA 600
 CTGCGAGAAA CTGCAATCTG TTTGCTCAGA CATTTCCTCA CACATTGATG AAACCTACCT 660
 60 GATGTTCTGG ATCGGGGTCA CCAGCGTACT GCTTCTGTTC ATCGTGTATG CGTACATGTA 720
 TATTCTCTGG AAGCTCTACA GCCACGCCBT CCGCATGATT CAGCGTGGCA CCCAGAAGAG 780
 CATCATCTC CACACGCTCG AGGATGGGAA GGTACAGGTG ACCCGGCCAG ACCAAGCCCE 840
 CATGGACATT AGGTAGCCA AGACCCCTGG CTGTATCTTG GTGGGTGTGA TCATCTGCTG 900
 GGGCCCTCTG CTTCGAATCA TGGTGTATGA TGTCTTGGG AAGATGAACA AGCTCATTA 960
 65 GACGGTGTIT GCATCTGCA GTATGCTCTG OCTGCTGAAC TCCACGCTGA ACCCATCAT 1020
 CTATCTCTG AGGAGTAAGG ACCTGCGACA CGCTTTCCGG AGCATGTTC CCTCTGTGA 1080
 AGGCACTGCG CAGCTCTGCG ATACAGCAT GGGGACTCG GACTGCTGCG ACAACAAGC 1140
 AAACATGCA GCCAGTGTTC ACAGGGCGCG AGAAGCTGC ATCAGAGCA CGGTCAAGAT 1200
 TGCCAAGGTA ACCATGTCTG TGTCCACAGK CAGTCTGCC GAGGCTCTGT GA

Seq ID NO: 240 Protein sequence
Protein Accession #: NP_149421.1

70 1 11 21 31 41 51
 75 MALQIPSPAP SPLTGTWNAQ MTFSTKTSKE ENEENIQCGE NFMDIECFMV LNPSQQLAIAV 60
 LSLTLGTFIV LENLVLVLCV LHSRLRCR SYHFIGSLA ADLLGSVIFV YSPIDFHV 120
 RKDSRNVL FKLGGVTASL ASVGSLEFITA IDRYISIRP LAYKRIVTR KAVVAFCLM 180
 80 TIALVAVL PLEGNWCEKL QVCSNDFPH IDETYLMFWI GVTSLVLLFI VYAYMYLWK 240
 HSBVRMIRI GTQKSIH TSDEGKVQVT RPDQARMDIR AKTLVLLIV VLIICWGPLL 300
 INVYDFGKM MNKLKIVFA FCSMLCLNST VNPPIIYALR SKDLRHAFRS MFPSCBETAQ 360
 LDNSMGSDD CLKHANMAA VERAESCIK STVKIAKVT MSVSTDTSAE L

Seq ID NO: 241 DNA sequence
Nucleic Acid Accession #: NM_003596.1

Coding sequence: 82..1194

1	11	21	31	41	51	
5	GTAGACTGTC	CATGGCCTGA	ACATTTCCTG	AAAATCATT	TGAGCAAAAT	ATCTGTTTAA 60
	TAACAAGATA	ACCACATCAA	GATGGTTGGA	AAGCTGAAGC	AGAACCTACT	ATTGGCATGT 120
	CTGGTGATTA	GTTCCTGTAC	TGTGTTTAC	CTGGGCCAGC	ATGCCATGGA	ATGCCATCAC 180
	CGGATAGAGG	AACGTAGCCA	GCCAGTCAAA	TTGGAGAGCA	CAAGGACCAC	TGTGAGAACT 240
10	GGCCTGGACC	TCAAGGCCAA	CAAAACCTTT	GCCTATCACA	AAGATATGCC	TTTAAATATT 300
	ATTGGRCGGG	TGCTCTGGAG	TGCAACCA	CTCATGAGGG	CCATGCTGGA	CGCACATCCT 360
	GACATTGCT	GTGGAGAGGA	AACCAGGGTC	ATCCCCGAA	TCCTGGCCCT	GAAGCAGATG 420
	TGGTCACGGT	CAAGTAAAGA	GAAGATCCG	CTGGATGAGG	CTGGTGTTC	TGATGAAGTG 480
	CTGGATTCTG	CCATGCRAGC	CTTCTTACTA	GAAATTATCG	TTAAGCATGG	GGAGCCAGCC 540
15	CCTTATTAT	GTAAATAAGA	TCCTTTTGGC	CTGAAATCTT	TAACTTACCT	TTCTAGGTTA 600
	TTCCCCAAATG	CCAAATTTCT	CCTGATGGTC	CGAGATGGCC	GGGCATCAGT	ACATTCATAT 660
	ATTTCCTGAA	AAGTTACTAT	AGCTGGATTT	GATCTGAACA	GCTATAGGGA	CTGTTTGACA 720
	AAGTGAATC	GTGCTATAGA	GACCATGTAT	AACCAATGTA	TGGAGGTTGG	TTATAAAAG 780
	TGCATGTTGG	TTCACTATGA	ACAACTTGTC	TTACATCCTG	AACGGTGGAT	GAGAACACTC 840
20	TTAAAGTTCC	TCCAGATTCC	ATGGAACCA	TCAGTATGTC	ACCATGAAGA	GATGATTGGG 900
	AAAGCTGGGG	GAGTGTCTCT	GTCAAAAGTG	GAGAGATCTA	CAGACCAAGT	AATCAAGCCA 960
	GTCAATGTAG	GAGCTCTATC	AAAATGGGTT	GGGAAGATAC	CGCCAGATGT	TTTACAAGAC 1020
	ATGGCAGTGA	GTGCTCTAT	GCTTGCCAG	CTTGGATATG	ACCCATATGC	CAACCCACCT 1080
	AACCTAGTAA	AACCTGATCC	CAAAATTAT	GAAAACACTC	GAGGGTCTTA	TAAGGGAGAA 1140
25	TTCCAACCTAC	CTGACTTTCT	TAAAGAAAA	CCACAGACTG	AGCAAGTGA	GTAGCAGAAC 1200
	CAGGAGCTC	TTCCATACAT	GAGGAAAGAT	TGCTGCCCTT	TCAGCAGAAG	GGAATTTCT 1260
	AGGATTTGGCT	GTCCCTGCTC	AAGCTTGGTG	GAGCGTCTGC	ACCTTGGCTG	CGCCGCTCTG 1320
	GCATTTGCCA	GTTCCTCTCC	ACTGAGAGGA	TGGAGGTGTC	CGCACAGCTT	TGGGCTCTGT 1380
	GAGGGATCTG	CCTCCTGAGC	AAAGAGCTCT	TGATCCCGAT	TTCATGCACA	GGCCTGCGAT 1440
30	AAGGAGCCCA	GAGGAAACAT	GTGTTTCTCT	TTAAACCTCC	TCTTGTCTCT	TTTCTTATCA 1500
	TTATGACGTT	TGTTTTCAAG	GAGAGGGTTT	AAAAATGGGA	TCTGTGAAGC	AGACTTGGCG 1560
	AGTCTCTCTT	TGAATAGGT	TGCTGTGACA	TGTTCTAATG	TTTTGTAGAA	CACGTGTGCC 1620
	TGTTTAAGTG	TATTGATGTG	AATAATATTA	AATATCTTAA	TTATTTAATT	CATTGTATTT 1680
35	TTTCTGAGAA	GTGGGAAAT	TACCATTA	CATTTACAAC	CTAATGACCT	TGTATTTTAA 1740
	TTTTTCAAAA	TAAAGCTTT	CAATGTGA			

Seq ID NO: 242 Protein sequence
Protein Accession #: NP_003587.1

1	11	21	31	41	51	
40	MVGLKQMLL	LACLVISSVT	VFYLQHAME	CHRIEERSQ	FKLESTRTT	VRTGLDLKAN 60
	KTFAYHKMP	LIFIGVPRB	GTLMRAMID	AHPDIRGEE	TRVPRILAL	KQWRSRSKE 120
	KIRLDEAGVT	DEVLDSAMQA	FLLEIIVKHG	BPAPYLONKD	PFALKSLTYL	SKLFFNAKFL 180
45	LMVRDGRASV	HSMISRKVTI	AGFDLMSYRD	CLTKWNRABE	TMYNQCMFVG	YKCKMLVHYE 240
	QLVLHPRNM	RLLKFLQIP	WNESVLHHEE	MIGKAGGVSL	SKVERSTDQV	IKFVNVGALS 300
	KWVGKIPDV	LQEMAVIAPM	LAKLGYDPYA	NPFNYGKDPD	KIENETRRVY	KGEFQLPDLF 360
	KERQTEQVE					

Seq ID NO: 243 DNA sequence
Nucleic Acid Accession #: NM_001492.3
Coding sequence: 1395..2513

1	11	21	31	41	51	
55	ACGCGGGGCG	CGCGGCTCCG	TCGGCTACCG	CGGGCGGGCG	CAGGCGACCG	GCACGGCGGG 60
	CGAGCGGGCG	GTATGGCGGC	GGCGGGGGCC	CGCGCGGGGC	CGAGCGGGGC	CGAGCGGGGC 120
	CGAGCTACG	CGACCTAGT	CGAGCGGGCG	TGGGGCAGCG	CGCTGGCGCG	GGCGCGGGGC 180
	TGCAGGACT	CGGGCTGGGG	GCTGGCGGCT	CGCGGGCTGG	CTGAGCAAGC	GCACCTGGCG 240
60	CGCGCGGGCG	TGCTGGCTGG	GGCGCTGGCG	CGCGGGGGCT	GGAGCGGGCT	GGCGCTGGCG 300
	GGCACTGGCG	GGCTCTTTGG	GGCGCTGGCG	AAGCGGTGCT	GGCTCCAGCC	CAGAGATGCC 360
	GGCAAGATGC	CGAGAGCGCG	TGGGAAGTTT	CTCTCTACCC	TGGGCGAGCT	GAGCTACAGT 420
	GGCTACCTGC	TGTTTGGCAC	CGACTACCCC	TTCTTCCATG	ACCCACCATC	TGTCTTCTAC 480
	GACTGGAGCG	CGGGCATGGC	AGTGCCACCG	GACATTGCAG	CCGCTTACCT	GCTCCAGGGA 540
65	AGCTTCTATG	GGCACTCCAT	CTACGCTAGC	CTATACATGG	ACACCTGGCG	CAAGGACTCG 600
	GTGGTCAATG	TGCTCCACCA	GGTGGTCACT	CTCATCTCTA	TGCTCTCTCT	CTACGCTCTC 660
	CGGTACACCA	ATGTGGGCAT	CCTTGTGCTC	TCTCTGCAAG	ATATCAGTGA	CGTGCAGCTT 720
	GAGTTCAACA	AGCTCAACAT	TTACTTCAAG	TCCCGCGGCG	GCTCTTACCA	TCGGCTGCAT 780
	GGCTTGGCAG	CAGACTTGGG	CTGGCTCAGC	TTGGCTTCA	GCTGGTCTCT	GTTTGGCTCT 840
70	TACTGGTTCC	CGCTCAAGGT	CCTGTATGCC	ACCAGTCACT	GCAGTCTGGG	CACGCTGGCT 900
	GACATCGGCT	TCTACTCTTT	CTTCAATGCG	CTCTGCTGCT	TGCTTACCTT	TATGAACCTC 960
	TACTGGTTCC	TGACATCGGT	GGCGTTTCCA	GCCAAAGTGT	TGACAGGCCA	GGTGCAAGAG 1020
	CTGAGGAGCC	TGGGGGAGTA	TGACACAGCC	GAGGCGGAGA	GCTTGAAGCC	CAGCAAGGCC 1080
	GAGAGGCCAC	TGAGGAACGG	CCTGGTGAAG	GACAAGCGCT	TCTGAACCTC	TGGGCGGGCG 1140
75	CCCGCTGGAC	CGGGCGGGCG	CCCGAATACC	CGGGCGGGCG	TCCCGTCTCT	TGGGCGGGCG 1200
	TCCACCGGCT	CTCACTCTGC	TGCTCTAGGG	CGGGCGGGCG	CTCCCGTGGG	ACCGCGGGCG 1260
	CTCATCTGCG	CTCATCTTCC	CGGGCGGGCG	CGGGCGGGCG	CTTGGCGGCT	CGGGGACCTC 1320
	GGCGCGGGCG	TCAGGCGGCT	GGTCCCGGGC	CGGGCGGGCG	CTTGGCGGCT	CTTCTGCTAT 1380
	CGCTTGGGAG	BAGAGTGCCA	CGGGCGGGCG	AGGGTCCCTG	CGGGCGGGCG	CTTCTCTCTC 1440
80	TCTTGGCGCT	GCTGCTGGCC	TGGCTGGCCG	TGACCGGGCG	CCCGGTGGCT	CCAGGCGGGC 1500
	CGCGCGGGCG	CTTCAAGGCT	CTAGGACTGC	GCGATGAGCC	CCAGGGTGGC	CCAGGCTCTC 1560
	GGCGGGTTCC	CCCGGTCTAG	TGGCGGCTGT	TTCGAGCGCG	GGAGCGGGCG	GAGACAGGCT 1620
	CTGGCTGGCG	CGGGCGGGCG	CCCTGCAACC	GTGCAAGCTG	GAGGAGCTGG	GAGGAGCTGG 1680
	GGGTGGCGGG	AAACATCGTG	CGGCACATCC	CGGAGCGGGC	TGGCGGGCGC	CGGGCGGGCG 1740
	AGCGTGTCTC	GGCGCGGGCG	CATTGGCGCT	AGTGAGACGT	CGTCTTGCAC	CTGTGGCGCT 1800

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

TGGAACCCGC TGAGCGCCCG AGCCGGGCCG GCCTGGAGCT GCCTTTCGCG GCGGCGGCGG 1860
 CCGAGCCCCG GGAGGGCGGC TGGGAGCTGA GCGTGGCGCA AGCGGGCCAG GGCBCGGGCG 1920
 CCGAACCCCG GCGGTGCTCG CTCGCCAGT TGGTCCCGC CCGGGGCGG CCACTGCGCG 1980
 CCGAGCTGCT GGGCGCCGCT TGGGCTCGCA ACGCTTCATG GCGCGCGAGC CTCGCGCTCG 2040
 CCGTGGCGCT ACGCCCCCGG GCGCTGCGC CCGCGCGCG GCTGGCCGAG GCGCTGCTGC 2100
 TGCTGGTGAC CCTGAGCCCG CGCTGTGCCC ACCCCCTGCG CCGCGCGCGG CGCGAGCGCG 2160
 AACCCGTGTT GGGCGCGCGC CCGGGGGGCG CTGTCGCGC GCGCGCGCTG TACGTGAGCT 2220
 TCCGCGAGGT GGGCTGGCAC CGCTGGGTCA TCGCGCGCG CCGCTTCCTG GCCAACTACT 2280
 GCCAGGCTCA GTGCGCGCTG CCGCTGCGCG TGTGCGGGTC CCGGGGGCGG CCGCGCTCA 2340
 ACCAGCTGT GCTGCGCGCG CTCATGCACG CCGCGCGCGG GCGAGCGCGC GACCTGCCCT 2400
 GCTGCGTGCC CCGCGCGCTG TCGCCATCT CCGTGTCTCT CTTTGACAAC AGCGACAACG 2460
 TGTGCTGCG CGATGATGAG GACATGGTGG TGGACGAGTG CCGCTGCGCG TAACCCGGGG 2520
 CCGGCGAGGA CCGGGGCCCA ACAATAAATG CCGCGTGGTC TGCTC

Seq ID NO: 244 Protein sequence
 Protein Accession #: NP_001483.2

1 11 21 31 41 51
 20 MPPFPQPCG HLLLLLLALL LPSLPLTRAP VPPGPAALIL QALGLRDEPQ GAPRLRFVPP 60
 VMWRLFRRRD FQSTRSGSRR TSPGVTLQPC HVEELGVAGN IVRHLPDRCA PTRASEPVSA 120
 AGHCPEWTVV FDLSAVEPAB RPSRARLELR FAAAAAAPF GGWELSVQA QGAGADPGP 180
 VLLRLQVLPAL GPPVRAELLG AAWARNASWP RSLRLALALR PRAPAACARL AEAAILLVTL 240
 25 DPLRLCHPLAR PRDRAEPVLG GPGGACRAR RLYVSPREVQ WRRWVIAPRG FLANYCQGGC 300
 ALPVALSGSG GPPALNHAUL RALMHAAAPG AADLPCCVPA RLEPISVLFF DNSDNVVLRL 360
 YEDMVVDECG CR

Seq ID NO: 245 DNA sequence
 Nucleic Acid Accession #: NM_021267.1
 Coding sequence: 17..1125

1 11 21 31 41 51
 35 ACGCGGGGCG CCGGCTCCG TCGGCTACCG CGGGCGGGCG CAGGCGACGG GCACGGCGGG 60
 CGAGCGGGCG GTATGGCGGC GCGGGGCGCC GCGGCGGGGC CGAGCGGGCC GAGGCCATG 120
 CCGAGCTACG CCGAGCTAGT GCAGCGCGGC TGGGCGAGCG CCGTGGCGCG GCGCGCGGGC 180
 TGACGAGACT GCGGCTGGGG GCTGGGCGGT CCGGCGCTGG CTGAGCACGC GCACCTGGCG 240
 CCGCCGAGC TGTGCTGCT GCGGCTGGGC GCGCTGGGCT GGACCGCGCT GCGCTCGCG 300
 40 GCGACTGCGC GCGCTCTTG GCGGCTGGCG AAGCGGTGCT GCGTCCAGCC CAGAGATGCC 360
 GCGAGAGTGC CCGAGAGCGC TTGGAAGTTT CTCCTTACCC TGGGCGAGTG GAGCTACAGT 420
 GCGTAOCTGC TGTTTGGCAC CGACTACCCC TTCTTCCATG ACCCACCATC TGTCTTCTAC 480
 GACTGGACGC CCGGCTAGGC AGTGCCACGG GACATTGCAG CCGCTTACCT GCTCCAGGGA 540
 AGCTTCTATG GCGACTCCAT CTACGCTACG CTATACATGG ACACCTGGCG CAGGACTCG 600
 45 GTGGTCTATG TGTCCACCA CGTGGTCACT CTCATCTCCA TCGTCTCTTC CTACGCTTTC 660
 CCGTACCA CA TGCTGGGCTC CTTGTGCTC TTCTGCAAG ATATCAGTGA CGTGCAGCTT 720
 GAGTTCACCA AGCTCAACAT TTACITCAAG TCCCGCGGCG GCTCCTACCA TGGGCTGCAT 780
 GCGTGGGCG CAGACTTGGG CTGCTTCAGC TTGGGCTTCA GCTGGTCTG GTTCGCTCTC 840
 TACTGCTTCC GCGTCAAGGT CCGTATGCG ACCAGTCACT CGAGTCTGCG CAGCGTGCCT 900
 50 GACATCCCTT TCTACTCTT CTTCATGCG CTCCTGCTGC TGCTCACCTT TATGAACCTC 960
 TACTGCTTCC TGTCATCTGT GCGTCTTGA GCGAGGTGT TGACAGGCCA GGTGCAAGAG 1020
 CTGAAGGACC TCGCGGAGTA TGACACAGCC GAGGCGCCAG GCGTGAAGCC CAGCAAGCC 1080
 GAGAAGCCAC TGAGGAACCG CCGTGGTAAG GACAGCGCT TCTGAACCCC TGGGCGCGCG 1140
 CCGCGTGCAG CCGGCGCCAC CCGGAATACC CCGGCAACGC TCCCGTCTCT TGGGCGCGCG 1200
 55 TCCACCCCTC CCACTCTGCG TCGTCTAGGG CCGCGCCAC CTCCCTTGGG ACGCGCGCC 1260
 CTCATCTGCG CTCGATTTCC CCGCCAGGCC CCGCAGGCC CCGCGCGCTC CCGGAGACCC 1320
 GCGCCCGCCC TCAGCCCACT GGTCCCGGCG CCGCGCGGAC CCGTGGCACT CTCTGGTCT 1380
 GCGCTGGGAG GAAGATGCCA CCGCGCGAGC AAGGTCCCTG CGGCGCACAC CTCTCTCTCC 1440
 TCGTGGGCTC GCGTCTGCG TCGTCTGCG TGAACCGCG CCGCGTGGCC CAGGCGCCAG 1500
 CCGCGCGCTC GCTCCAGCT CTAGGACTGC GCGATGAGCC CCGGCGGCG CCGAGGCTCC 1560
 60 GCGCGCTTCC CCGGCTCATG TGGGCGCTGT TTGAGCGCG GAGCCCGCAG GAGACAGGT 1620
 CTGCTGCGG CCGGAGCTTC CCGGCGTCA CCGTCAACC GAGCCACGTG GAGGAGCTGG 1680
 GGGTGGCGCG AAACATCGTG CCGCACATCC CCGACCGCG TCGCGCCACC CCGGCTCTCG 1740
 AGCTGTCTC GCGCGCGCG CATTCGCTG AGTGACAGT GGTCTTGGAC CTGTCGGCTG 1800
 65 TGGAAACCGC TGAGCGCCCG AGCGGGGCC GCGTGGAGCT GCGTTTCCCG GCGGCGCGCG 1860
 CCGGAGCCCG GAGGGCGGCG TGGAGCTGA GCGTGGCGCA AGCGGGCCAG GCGCGGGCG 1920
 CCGACCCCGG CCGGCTGCTG CTCGCGCAT TGGTGGCGCG CCGGGGCGG CCGTGGCGCG 1980
 CGGAGCTGCT GCGCGCGCT TGGGCTGCGA ACGCTCATG CCGCGCGAGC CTCGCGCTGG 2040
 CGTGGCGCT ACBCCCCCG GCGCTGCGG CCGCGCGCG CCGGCGCGAG GCGTGGCTGC 2100
 70 TGTGCTGAC CCGGACCGG CCGCTGTGCC ACCCGCTGCG CCGCGCGCG CCGGACCGCG 2160
 AACCGTGTG GCGCGCGGCG CCGGGGGGCG CTGTCGCGC GCGGCGGCTG TACGTGAGCT 2220
 TCCGCGAGGT GGGCTGGCAC CGCTGGGTCA TCGCGCGCG GCGCTTCTCG GCCAACTACT 2280
 GCGAGGCTCA GTGCGCGCTG CCGTGGCGC TGTGCGGCT CCGGGGGCG CCGGCGCTCA 2340
 ACCAGCTGT GCTGGCGCG CTCATGCACG CCGCGCGCG GCGAGCGCG GAGCTGCCCT 2400
 75 GCTGCTGCC CCGCGCGCTG TCGCCATCT CCGTGTCTCT CTTTGACAAC AGCGACAACG 2460
 TGTGCTGCG CGATGATGAG GACATGGTGG TGGACGAGTG CCGCTGCGCG TAACCCGGGG

Seq ID NO: 246 Protein sequence
 Protein Accession #: NP_067090.1

1 11 21 31 41 51
 80 MAAAGPAAGP TGPEPMPSTA QLVQRWGS A LAAARGCTDC GWGLARRGLA EHAHLAPPEL 60
 LLLALGALGW TALRSAATAR LFRPLAKRCC LQPRDAKMP EBAWKFLFYL GSWSYAYLL 120
 FGTIYPPFHD PPSVFDWTF GMAVERDIAA AYLLQGSFTG HSIYATLYMD TWKDSVVML 180

LHHVTLILI VSSYAFRYHN VOILVFLHD ISDVQLEFTK LNIYFKSRGG SYHRLHALAA 240
 DLGLCSFGFS WFWRLYWF LKVLVATSHC SLRTVPDIPP YFFFNALLIL LTLMLNYWFL 300
 YIVAPAAKVL TGQVHEKDL REYDTAEQSL KPSKAKEPL RNGLVVKDRF

5

Seq ID NO: 247 DNA sequence
 Nucleic Acid Accession #: NM_002081.1
 Coding sequence: 222..1898

10 1 11 21 31 41 51
 | | | | | |
 GGCTGCCCGA GCGAGCGTTC GGACCTCGCA CCCCGCGCGC CCCGCGCGCG CGCGCGCGCC 60
 GGCTTTTGT GTCTCGCCCT CCGCGCGCGC CGCGCGCTCT GGACCGCGAG CGCGCGCGCG 120
 CGGACCTTG GCTCTGCCCT TCGCGGCGCG GAACTGCGCA GGACCGCGCC AGGATCCGAG 180
 AGAGGCGCGG CGCGGTGCCG GGGGCGCGCG CGCGCGCGCG CATGGAGCTC CGCGCGCGAG 240
 15 GCTGGTGCT GCTATGTGCG GCGCGAGCGC TGGTCCGCTG CGCGCGCGCG GAACCGCGCA 300
 GCAAGAGCGG GAGCTCGCGC GAGGTCCGCG AGATCTACGG AGCCAGAGGC TTCAGCTTGA 360
 GCGAGCTGCC CCAGGCGGAG ATCTCGGGTG AGCACTCGCG GATCTGTCCC CAGGGCTACA 420
 CCGCTGTGAC CAGCGAGATG GAGGAGAACC TGGCCAAACG CAGCCATGCC GAGCTGGAGA 480
 CGCGCTCCCG GGACAGCAGC CGCGTCTGCG AGGCCATGCT TGCCACCCAG CTGGCGAGCT 540
 20 TCGATGACCA CTTCAGCAC CTGCTGAACH ACTCGAGCG GACGCTGCAG GCCACCTTCC 600
 CGCGCGCTCT CGGAGAGCTG TACACGCAGA ACGCGAGGCG CTTCGCGGAC CTGTACTCAG 660
 AGCTCGCGCT GTACTACCGC GGTGCCAACC TGCACTTGA GAGAGCGCTG GCGGAGTTCT 720
 GGGCCGCTCT CTCTGAGCGC CTCTTCAAGC AGCTGCACCC CCAGCTGCTG CTGCTGATG 780
 ACTACTCGGA CTGCTCGGCG AAGCAGGCGG AGGCGCTGCG GCGCTTCCGG GAGGCGCGCA 840
 25 GAGAGCTGCG CCGTGGGCGC ACCCGTGCCT TCGTGGCTGC TCGCTGCTTT GTGAGGCGCC 900
 TGGGCGTGGC CAGCGACGTC GTCCGGAAAG TGGCTCAGGT CCCCTGCGC CGGAGTGCT 960
 CGAGAGCTGT CATGAGCTG GTCTACTGTC CTCACTGCTT GGGAGTCCCC GCGCGCAGGC 1020
 CCGTCCCTCT TAATGCGGA AATGTGCTCA AGGCTGCTT TCCCAACAG GCGGACCTGG 1080
 ACGCGAGTGT GAGGAACCTC CTGACTTCCA TGGTGTCTAT CACCGACAAG TTCTGGGTGA 1140
 30 CATCGGCTGT GAGAGTGTG ATCGGCGAGG TGCACAGCTG GCTGGCGGAG GGCATCAAG 1200
 CCTCCAGGA CAACAGGAGC ACGCTCAGGG CCAAGGTCTT CCAGGCTGCG GGAACCCCA 1260
 AGGTCAACCC CCAGGCGCTT GGGCTGAGG AGAAGCGCGC CCGGCGCAAG CTGGCGCGCG 1320
 GGGAGAGGCG ACCCTCAGGC ACCTGGAAGA AGCTGCTCTC TGAAGCCAA GCGGAGCTCC 1380
 GCGAGCTCCA GAGCTTCTGG ATCAGCTTCC CAGGAGCACT GTGAGTGAG AAGATGCGCC 1440
 35 TGAGCACTGC CAGTGATGAC CGCTGCTGGA ACGGATGCGC CAGAGGCGCG TACCTCCCG 1500
 AGGTCTGCGG TGACGCGCTG GCCAACAGGA TCAACAACCC CAGGTGAGG GTGGACATCA 1560
 CCAAGCCGGA CATGACCTC CGCGAGCAGA TCACTGAGCT GAAGATCATG ACCAACCGCG 1620
 TGGCAGCGCG CTACACGCGC AACGAGCTGG ACTTCAGAGA CGCCAGTGAC GACGCGAGCG 1680
 GCTCGGCGAG CGGTGATGCG TGTCTGGATG ACCTCTGCGG CCGGAAGGTC AGCAGGAAGA 1740
 40 GCTCCAGCTC CGGAGCGGCC TTGACCCATG CCTCCGAGG CCGTCTCAGG CAGGAAGGAC 1800
 AGAAGAGCTC GGTGCGCAGC TGCCCGGAGC CCCGAGCTT CCGTCTGCGC CTCTCTCTCT 1860
 TCTTGGCGCT TACAGTAGCC AGGCGCGCGT GGGGTAACT GCGCCAGGC CCGAGGAGCA 1920
 GAGGCGCAAG ACTGACTTTC CCAAAATAC AACACAGAG ATATTTAATT CACCTCAGCC 1980
 TGGAGAGGCG TGGGTGCGGA CAGGAGGCGC CGCGCGCTCT GAGCAGGCGC AGGCGCAGAG 2040
 45 GTCCAGCGCC CAGGCTGCGC CTGCGCTGCG TTTCTGCGCT TTAATTTTGT ATGAGTCTCT 2100
 CAGGTCACTT GGGAGCCAGT GTGCCCAAAA GCCATGTATT TCAGGAGCTT CAGGCGCACC 2160
 TCGGCTGCGC TAGCCTTCCC CCGAGCTCCC TGCAACCGCG CAGAGCAGC CCTCGAGGC 2220
 CTACAGAGGA GGCCTCAAG CAACCGCGTG GAGCCACAG CGAGCCTGTC CCTTCTTCCC 2280
 CGCTCTCTTC CACTGGGACT CCGAGCAGAG CCGACCAAGC AGCCTTGGCC CACCCCGGCG 2340
 50 CCTCCAGAGA AGCCCGGCGC GGGCTGTCTG GGTGTCCGCG ATCCAGGGTC TGGCAGAGCC 2400
 TCTGAGATGA TGCATGATGC CCTCCCTTCA GCGGAGGCTG CAGAGCCCGG CCGCAGCTCC 2460
 CTGCGCGCTT GAGGCGCGCC AGCGTCTGCA GGGTGAAGCC TGAGACAGCA CCAGTCTTGA 2520
 GGAATCTGAG GACTGTCTTC CCACAGACCC TGCAGTGAGG GGCCTTCCAT GCGCAGATGA 2580
 GGGGCGACTG ACCCACTGCG GCTTCTGCTG GAGGAGGAGA AGCTGGGCCC AAAGCGCCAG 2640
 55 GAGGCGAGCG TGGCTCTGCG CAGTGTGGGC TGCCCTTCCG ACACAGGGCT CAGGCGCAG 2700
 GCGTGTCTGG GGTCCAGGGC TGTGAGAGGA CCGCGAGGCG TGAGGAGCAG CCAGGACCCG 2760
 CCTGTCTCCA TCTTCACTCA GATCAGGAAC CAGGCGCTCC CTGTTCAAGG TGACACAGGT 2820
 CAGGCTCAGG GCGTGTCTCC GGTGTCTACC TGTCTCAGAG GATGCTGCTG GCTGTGAGA 2880
 CCGCGACTCG CACACGGGAA TGCTTAGGTC CCTTCCCGAC CCGAGCAGCT GCACTGAGCA 2940
 60 GCAAGGCGAC CTGATAGTT AAGGGCTTTT CCAAACTGCG ATCCATTAC TGACACTTCC 3000
 TGTCTTGTGT CATGAGAGAG TGTTCGCTCC TCCAGATGG CTTCGAGGC CCGCAGGCGC 3060
 CACCTTGGAC CCGTGTGACC TCTGTGACT CACTGAGGCC ATCAGGCGCC TGCCCGAGGC 3120
 CTGAGCGGCG CCTCTTCCC TCTGTGCGC CAGCTGCCAG GTGGCGCTGG GAGGCGGTGG 3180
 TGTGTGCTTG GGAAGGGGTC CTGAGGGGG AGGAGGACTT GAGGCGCTG GGGGCGAGCT 3240
 65 TCTGAGACCG ACTGACCTG AGGAGGCGCG TTAGTGCTGC TTTGCTTTTC ATCAGGCTCC 3300
 CGCAGAGTGG ACAGAGGTCC CCGGTTGCTG GTCAGGTCCC CATGGCTTGT TCTCTGGAAC 3360
 CTGACTTTAG ATGTTTGGG ATCAGGAGCC CCGACACAG GCAAGTCCAC CCATAATTA 3420
 CCTGCGAGT GCGAGGGTGG GCTGGGGACT CTGGCAGAGT GATGCCGCGC GCGAGGACAG 3480
 CAGCACTCCC GCTGCACACA GACGGCTAG GGGTGGCGCT CAGACCCAC CCTAGCTCA 3540
 70 TCTCTGGAAG GGGCAGGCTT GAGTGTCTAC TGGTCAGGCG AGTGGCCAGG CCGTGTGCT 3600
 CCTTCTTCCA CAGGTCCCG CCAAGGCTCA GTGTCAAGCG GTGACGCTG TCTTTTGTAG 3660
 TCTTGTATG AATAAAGGC TGGAAACCTA AA

75

Seq ID NO: 248 Protein sequence
 Protein Accession #: NP_002072.1

1 11 21 31 41 51
 | | | | | |
 80 MELRARGNWL LCAAAALVAC ARGDPASKSE SCEVRQIYG AKGFSLSDVP QAEISGEHLR 60
 ICQGGTCTCT SEMENLANR SHAELETALR DSRVLIQAML ATQLRSFDH FQHLNDSER 120
 TLQATPFGAF GELYTQNARA FRDLYSELRL YVRGANLHLE ETLAEFWARL LERLPQLHP 180
 QLLLPDDYLD CLGKQAEALR PFGAPRELR LRATRAFVAA RSFVQGLGVA SDVVRKVAQV 240
 PLGFECSRVA MKLVYCAHCL GVPGARPCPD YCRNVLKGLC ANQADLDAEW RNLLDSMVL 300
 TDKFWGTSGV ESVIGSVHTW LAEALNALQD NRDTLTAKVI QGCGNPKVNP QGPGPERKRR 360

RGKLAPRRER PSGLTLEKLVS KAKAQLRDVQ DFWISLPQTL CSEKMALSTA SDDRCWNGMA 420
 RGRYLPVVMG DGLANQINNP EVEVDITKPD MTRQQIMQL KIMTNRLRSA YNGNDVDFQD 480
 ASDDGSGSGS

5

Seq ID NO: 249 DNA sequence
 Nucleic Acid Accession #: NM_001492.3
 Coding sequence: B..1864

10 1 11 21 31 41 51
 | | | | | |
 GAAGGCCATG GTCTCCCCAC GGATGTCCGG GCTCCTCTCC CAGACTGTGA TCCTAGCGCT 60
 CATTTTCCTC CCCACAGACAC GGCCCGCTGG CGTCTTCGAG CTGCAGATCC ACTCTTTCGG 120
 GCGGGGTCCA GGCCCTGGGG CCCCGCGGTC CCCCCTGCAGC GCCCGGCTCC CCTGCCGCCT 180
 15 CTCTCTCAGA GTCTGCCTGA AGCTTGGGCT CTCAGAGGAG GTCGCGGAGT CCGCGTGCCT 240
 CTTGGGCGGG GCGCTGAGCA CGCGGAGACC GGTCTACACC GAGCAGCCCG GAGCGCCCGC 300
 GCGTGTCTTC CCACTGACCG ACGGGCTCTT GCAGGTGCCC TTCGGGAGCG CTTGGCCTGG 360
 CACCTTCTCT TTCATCATCG AAACCTGGAG AGAGGAGTTA GGAAGACCAGA TTGGAGGGCC 420
 CGCTTGGAGC CTGCTGGGCG GCGTGGCTGG CAGCGGCGCG TTGGCAGCCG GAGGCCCGTG 480
 GGCCCGGGAC ATTCAAGCGG CAGGCGCCTG GAGCTGCGC TTCTCGTACC GCGCGCGCTG 540
 20 CBAGCGGCGT GCGCTGCGGA CGCGTGCAC GCGCCTCTGC CGTCCGCGCA GCGCCCGCTC 600
 GCGGTGCGGT CCGGAGTGC GCGCTGCGC ACCGCTCGAG GACGAATGTG AGCGCGCGCT 660
 GGTGTGCCBA GCAGGCTGCA GCCCTGAGCA TGGCTTCTGT GAACAGCCCG GTGAATGCCG 720
 ATGCTAGAGG GGCTGAGCTG GACCCCTCTG CAGGTTCCCT GTCTCCACCA GCAGCTGCCT 780
 CAGCCCGCAG GGCCCGTCTT CTGCTACCA CCGATGCTT GTCCCTGGGC CTGGGCGCTG 840
 25 TGACGGGAGC CGCTGAGCA ATGAGGCTAG CTGTAGTGAG ACACCCAGGT CCTTTGAATG 900
 CACCTGCCCG CCGGGTCTT ACGGGCTGCG GTGTGAGGTG AGCGGGGTGA CATGTGCAGA 960
 TGGACCTGCT TTCAACGCGG GCTGTGTGTG CGGGGTGCA GACCTGACT CTGCTACAT 1020
 CTGCACTGCG CCACCTGGTT TCCAAGGCTC CAACTGTGAG AAGAGGGTGG ACCGGTGCAG 1080
 CTTGCAAGCA TGCCGCAATG GCGGACTCTG CTTGGACCTG GGCCACGCCC TGCGCTGCCG 1140
 30 CTGCGCGGCG GCGCTGCGGG GTCTCTGCTG CAGACACGAC CTGAGCAGCT GCGCGGCGCG 1200
 CGCTCTGCTT AACCGCGGCA CGTGTGTGGA GCGCGCGCGC GCGCACGCTT GCTCTGCGC 1260
 GCTGGGCTTC GCGCGCGCGC ACTGCGCGCA GCGCGCGGAC CGTGTGCGCG CGCGCCCGTG 1320
 TGCTCADGBC GGCCGCTGCT ACGCCCACTT CTCGGGCTCT GTCTGCGCTT GCGCTCCCGG 1380
 35 CTACATGGGA GCGCGGTGTG AGTTCCGCTT GCACCCCGAC GCGCGCAGCG CCTTGCCCGC 1440
 GGCCCGCGCG GCGCTGAGCG CGCGGACGCC TCAGCGCTAC CTTTTCCTTC GCGCTCTGGG 1500
 ACTGCTCGTG GCGCGGCGCG TGCGCGCGCG TGCGCTCTTG CTGGTCCACG TGCGCGCGCG 1560
 TGGCCACTCC CAGGATGCTG GGTCTGCTTT GCTGGCTGGG ACCCGGAGC GGTCAATCCA 1620
 CGCACTCCCG GATGCACTCA ACAACCTAAG GACGCGAGG GGTTCGCGGG ATGGTCCGAG 1680
 40 CTCGTCGCTA GTGTGGAATC GCGCTGAAGA TGTAGACCTT CAAGGGATTT ATGTCTATAT 1740
 TGCTGCTGCT ATCTGAGCTC GGGAGGTAGC GACGCGCTCT TTCCCGCGCG TACACACTGG 1800
 GCGCGCTGGG CAGAGGCGAG ACCTGCTTTT TCCCTACCTT TCCTCATATC TGTCCGTGAA 1860
 ATGAATTGGG TAGAGTCTCT GGAAGGTTTT AAGCCCATTT TCAGTTCTAA CTTACTTTCA

45

Seq ID NO: 250 Protein sequence
 Protein Accession #: NP_058637.1

1 11 21 31 41 51
 | | | | | |
 50 MVSPRMGSL SQTIVLALIF LQTRPAGVF ELQIHSFPG PGPGAPRSPC SARLPCLRF 60
 RVCLKPLGE EAAESPCALG AALSARGPVY TEQPGAPAPD LPLFDGLLQV PFRDANPGTF 120
 SFILITWEEB LGQIGGPAW SLLARVAGRR RLAAGGPARW DIQAGANWEL RFSYRARCPE 180
 PAVGTACTRL CRPRSAPSRC GPGLRPFCAF EDECEAPLVC RAGCSPEHGF CBQPGECRCL 240
 BHWIGPLCTV PVSTSSCLSP RGPSSATTGC LVPGPGPCDG NPCANSGSCS ETPRSFECTC 300
 55 PRGFYGLRCE VSGVTCADGP CFNGGLCVGG ADPDSAYICH CPPGFQSNIC EKRVDRCBLQ 360
 PCRNGLCLD LGHALRCRCR AGFAGPRCEH DLEDCAGRAC ANGGTCVEGG GAHRCSCALG 420
 FGGRCRERA DFCARPCFAH GGRCAHFSG LVCAACPGYM GARCEPVEHP DGASALPAAP 480
 PGLRPGDPQR YLLEFALGLL VAAGVAGAAL LIVHVRRGH SQDAGSRLLA GTPEPSVHAL 540
 PDALMNLRTQ BSGGDGSSSS VDMNRPEVD PQGIYVISAP SIYAREVATP LPPPLETGRA 600
 GQRQHLLFPY PSSILSVK

60

Seq ID NO: 251 DNA sequence
 Nucleic Acid Accession #: CAT cluster

65 1 11 21 31 41 51
 | | | | | |
 GAAATATAAC CATTGCAATT AGAAAATATC CAAAATAGCC TGTATCTTTC CACGTGGCCT 60
 AGATTATTGA CAATCCCAAA TATACAATT TCTTTAAAA GTAGTACAAT TTCCTTTGTA 120
 GCTTCATATC CTTATATGAC TTCAGACTGG AGAAGCCTGT TAAACCACTG TTAGTTTCAG 180
 70 TTAGAAAGTC TGAGAGACTT TATACATAAA TCTCAATTT GGCTGTGTGA CAGTGCCAG 240
 AGTTTACTA CTGTAGTGAC CGTTGAGAAG ACCCTTGTIT ATTACATTT GAAGCACTGT 300
 TTGTGCAAC AACCTTTCAT TGTTAAGTGC CTGTATTCTT TTCATTACT TCATGTCCAG 360
 GGGTCTATT TACTAGAAC CATTGTCTAC TACAATTAAC ATTTACATTA CAAAGTGTGT 420
 GGTTTTCTTT TTCAAGGAGG TTCAATTAAG GCAATAAGAT GTTTGCTGGA GAAACCTATT 480
 75 GTTTACTGAA AGCACTCAAT GAAGTCAAT TACTGAAGCT TTTGCTTACA TCTTGTCTTT 540
 TTATGTAAAT ATGTTAAATA TAACATCTAA GGAATAATAA CAATATTATA ATTATGTGTT 600
 TGCCATTGTC ATATCAAACT TGCTTTGTAT CATACTAATG TTACATAACT TATCGATCAA 660
 TAAAAATACA TTTCAATGTT AAAAAAATA AAAAAAATA

80

Seq ID NO: 252 DNA sequence
 Nucleic Acid Accession #: Bos sequence

1 11 21 31 41 51
 | | | | | |
 AGGTACTGCC AGAAGAGGATC AGGACCTGGA GTCTGGCAAG AGGAAGACAG AGGCGCTGTGT 60

5
 10
 15
 20
 25
 30
 35
 40

```

GGGAAGCGAG TTGTTATCTT TGGTTATCTA GCTGTATGAG TGTATTGGTC TTCATAAAGC 120
TAGATAAACCG AAGTAAAAAA CTCTTCAAG ATCGCCGGGG AGCGTGTGAG AATGAAAGAC 180
TACAGCCBAG AGACAGTAAA AACCAGAAAG GTCCAGGAATA CTTATTGAAT CTAACCTTGT 240
TTTTGTTTGG TTTTTTTCCT TATGATTAAA GGTGGGATGA GAGAAAAATTA AATGACACAC 300
ACATGCTAAA ATATCAAGGT TCCAGATATG TCTTGAGAGG GGTGTGTGCA GCTGCAAAAG 360
AGAAGTGATAT AGTGATAATG AGTAAAGATG CATGTGCAGT TTGTTCTATT TTAAGGCAAA 420
AGTTATATCA GGGATTTTTT TCTTAGAAAG GTGTTGCAGA GATGTCGGT ACCTAGTTTA 480
AAAAATGATTC CATAATATGT AGACTTGGGC AGTTCCTTTG GGAGGCACCT CCCTCTCAAA 540
ATTTGAAGAT TGTGCTTGGG AATTACTTTA CATGTATTGT GGCTGTATGT CAATTGTATC 600
GAAATTAGAT TTTCAAGAAA GTTTATACTG GAAGGTTAAT AATTGTATAT TACTGAGGAC 660
TTAGAGCTAG CAGGCAAAAT GAAAAAATAA AAAAAAARG CAAGGGCTGA TTTTATTCT 720
TCTATTCAAA ATACAAGGAC AGATGCTTCT CTGTTCCAAG AGGTTTCCT TGAGGAAGCT 780
ACTGAAGCAG AAAGACATGA TGGAGACGAG ATCGCCTCCC CTCTTGTCAA AGTGTAAAA 840
AAATGTTCTG TCTTACTCTG CGCTTAGCAT TGGAAATGAA AGTGACATTI ACGCCACAC 900
CCAGTGTGAG TTTTCTCTCT TTTTGTGTTA AGGATGATCA GGTCTATCCA GGAAACAGCT 960
CTGGCATCCC AACTGTAAT AATTAGGACG TATATAGACC TGACAAAAAT GGAAAGGGGG 1020
TGGGAATCT GAGGTCTGT CTGCTCAAT TGATTCGCT AAACGGAATG CAGGAGATGT 1080
GAACGCGCAG ACCTCTCGAT TCCACGCTC GGGGCGAAGT GATAAAGCGG GCGCGGCGAG 1140
CCTATGACAG ACAGCCCTGT TGGGGGGTGG GGGTATGAAA AAAAAATCAA GTGCACACAC 1200
CATACTCATC TCCATCGCTT AAGAAAGTAA AGGCATTTC CACCCACAGC CATCTGCAGC 1260
TTCCCAATG CAGACACCAA CTGGTCTGTA GCTGTACAT AGTCTGCTCT TGTAAATTTT 1320
TTAACCCTG TTTAATCTGG CCATAATTAA GTTTGGCTTT CTTCGTTGTT TGAGATTTC 1380
AGAATTCAAG GCAAGCTAGT AGAAAGCAAT TCCAGAAAG TCCCATGACT GCGTGGCCCT 1440
AATGTCAAAA TCTCAGTCCA TGAGATTATG GCCTTGTGAC CACATTTTGT CTTTGTGTTT 1500
GGTGGGCGAA ATGTGTATAG AGATAAAATA CATATCTCTA TATAACAGTC GTTATTATTA 1560
TTTCATGAGG CTTTTTACCT CTTAACATGA TACATCTAGG AACTTGGTCT AATTGTGCTA 1620
GTAGATATCT GATGAGATTA AACTATAAA GTCATTAGCC TGTAGAAAGT TGATTATGAT 1680
AACAAATGA TAAAAAGTTT GTTTTGAAT AGTCTCAGCT AGATGGGCTT AATAGCCAT 1740
TTTAATGTAA TCTAAAAATA ACCTATGCCC TAGCAGAAAC TTTGGGCGCT TGGAGGTCCC 1800
CATTTGTGCC TTTTCTAAA AGTCTCTAAG TTTTCCATAT GTCAACCAAG AAACATTTGT 1860
ATAGGCATTT ACACAGAAAT ATAACCATG CAATTAGAAA ATATCCAAA TAGCCTGTTA 1920
TCTTCCAGCT GGCCTAGATT AITGACAATC CCAATATAC AATTTTCTT TAAAAGTAGT 1980
ACAAATTCCT TTGTAGCTTC AATTCCTTAT ATGACTTCAG ACTGGAGAAG CCTGTAAAC 2040
CACTGTTAGT TCTAGTTAGA AAGTCTGAGA GACTTTATAC ATAAATTTCT AATTGTGCTG 2100
CTGTACAGCT GTCAGAGTTT TACTACTGTA GTGACGGTGG AGAAGACCTT TGTTTATTTA 2160
CATTTGAAGC ACTGTTTGG CAAACACCTT TCTATGTTA AGTGCCTGTA TTCTTTCTAT 2220
TTACTTCTAG TCCAGGGGTG CTATTACCT AGAACCATTG TCTACTACAA TTAACATTTA 2280
CATTACAAAG TGTGTGGTTT TCTTTTCAA GGAGGTTCAA TTAAGGCAAT AAGATGTTTG 2340
CTGGAGAAAG CTATTGTTTA CTGAAAGCAC TCAATGAAGT CAAATTACTG AAGCTTTTGC 2400
CTACATCTTG GTCCTTTATG TAAATATGTT AATATAACA TCTAAGGAAA ATAAACATA 2460
TTATAATTAT GTGTTTGCCA TTGTCATATC AAACCTGCTT TGTATCATAC TAATGTTACA 2520
TAACCTATCG ATCAATAAAA ATACATTTCA AGTTT
  
```

45
 Seq ID NO: 253 DNA sequence
 Nucleic Acid Accession #: NM_001650.2
 Coding sequence: 40.1011

50
 55
 60
 65
 70
 75

```

1 11 21 31 41 51
| | | | |
GGGGCAGGCA ATGAGAGCTG CACTCTGGCT GGGGAAGGCA TGAGTGACAG ACCCAGAGCA 60
AGGCGGTGGG GTAAAGTGTGG ACCTTTGTGT ACCAGAGAGA ACATCATGGT GGCTTTCAAA 120
GGGGTCTGGA CTCGAAGCTTT CTGGAAGACA GTCAAGCCBG AATTTCTGGC CATGCTTATT 180
TTTGTCTTCC TCAGCCTGGG ATCCACCATC AACTGGGGTG GAACAGAAAA GCCTTTACCG 240
GTGACATAGG TTCTCATCTC CCTTIGCTTT GGACTCAGCA TTGCAACCAT GGTGCAGTGC 300
TTTGGGCTAG TTGAGGGTGG CCACATCAAC CCGTCAAGTA CTGTGGCCAT GGTGTGCACC 360
AGGAAGATCA GCATGCCCCA GTCTGTCTTC TACATCGCAG CCGAGTGCTT GGGGGCCATC 420
ATTGGAGCAG GAATCCTCTA TCTGGTCACA CCTCCAGTG TGGTGGGAGG CCGGGAGTGC 480
ACCATGTTTC ATGGAATCTI TACCGCTGGT CATGGTCTCC TGGTTGAGTT GATAATCACA 540
TTTCAATTGG TGTTACTAT CTTTGCCAGC TGTGATTTCA AACGGACTGA TGTCACTGGC 600
TCAATAGCTT TAGCAATTGG ATTTCTGTT GCAATTGGAC ATTTATTGTC AATCAATTAT 660
ACTGGTGCCA GCATGAATCC GCGCCGATCC TTTGGACCTG CAGTTATCAT GGGAAATTGG 720
GAAACCAATT GGATATATTG GGTGGGCC ATCATAGGAG CTGTCTCTGC TGGTGGCCTT 780
TATGAGTATG TCTTCTGTCC AGATGTTGAA TCAAAAGTTC GTTTTAAAGA AGCCTTCAGC 840
AAAGCTGCCC AGCAACAAA AGGAAGCTAC ATGGAGGTGG AGGACACAGG GAGTCAGGTA 900
GAGACGGATG ACCTGATTCT AAAACCTGGA GTGGTGATG TGATTGACGT TGAACGGGGA 960
GAGGAGAAAG AGGGGAAAGA CCAATCTGGA GAGGTATTGT CTTCAATATG ACTAGAAGAT 1020
CGCACTGAAA GCAGACAAGA CTCCCTAGAA CTGTCCTCAG ATTTCTCTCC ACCCATTAAG 1080
GAAACAGATT TGTATATAAT TAGAAATGTG CAGGTTTGTG GTTTCATGTC ATATTACTCA 1140
GTCTAAACAA TAAATATTTT ATAAATTTACA AAGGAGGAAC GGAAGAAACC TATTGTGAAT 1200
TCCAAATCTA AAAAAAGAAA TATTTTAAAG ATGTTCTTAA GCAATATATAT ACCTATTTTA 1260
TCTAGTTACC TTTCATTAAC AACCATTTT AACCCTGTGT CAAGATTGGG TTAAGTCTTG 1320
CCTCAGACGA CTCAAAGACA GGTCTATCAG CTTATTCCTT CTCTACTGGA ATATTGGTAT 1380
AGTCAATCTT TATTGAATA TTTATCTTAT TAACTGAGT TTAACATGG C
  
```

75
 Seq ID NO: 254 Protein sequence
 Protein Accession #: NP_001641.1

80

```

1 11 21 31 41 51
| | | | |
MSDRPTARRN GKCGPLCTRE NIMVAFKGVW TQAFWKAVTA EFLAMLIFVL LSLGSTINWG 60
GTEKEPFDVM VLISLCEFLS IATMVQCFGH ISGGHINPAV TVAMVCTREI SLAKSVFYIA 120
AQCLGAILGA GILYLVTPPS VVGGGLGVIMV RGNLTAGHGL LVELIITFQL VPTIFASCDS 180
KRTDVTGSLA LAIGFSAVIG ELFAINYTGA SMNPARSFGP AVIMGNWEMH NIYVWGPIIG 240
AVLAGGLYIEY VFCDPVEFKR RFKEAFSKAA QQTKGSYMEV EDNRSQVETD DLILKPGVVH 300
  
```

VIDVDGZEEK KGRDQSGEVL SSV

Seq ID NO: 255 DNA sequence
Nucleic Acid Accession #: U26742.1
Coding sequence: 325..1449

5
10
15
20
25
30
35
40

1	11	21	31	41	51	
CAGGAAACCC	TGGTACTGGC	AGCAGCCAGC	CTCTGCTGTG	CCCACATGAC	CCACAACCTCT	60
GGCAGCGGAC	CCGCGACTTC	CAACATTATT	AAATAATAAG	AAAGCGGCTC	CTACTCCAGG	120
CTCAAACTTC	CCTGCAGACC	AATGGACACC	TTCTAAGAGT	TTGGCGAGTC	AGTGACTGAA	180
GCSCCGCTCC	ATTCCAAGAT	AAATAGGATT	TACCAATCCT	TGGATGAAGT	GCTTGGGAAG	240
TCITTAAGTG	CCATAATCAA	CTGCCATTTT	AAAGAATATA	GATGGTTTGT	AAAAAGTTTAT	300
GCTGTCCCTT	CATTGAATTT	TAGAATGATT	GAAGATAGTG	GGAAAAGAGG	AAATACCATG	360
GCAGAAAGAA	GACAGCTGTT	TGCAGAGATG	AGGCTCAAG	ATCTGGATCG	CATCCGACTC	420
TCCACCTACA	GAACAGCATG	CAAGCTTAGG	TTTGTTCAGA	AGAAATGCAA	TTTGCACCTG	480
TGGGACATAT	GGATGTTCAT	AGAAGCATTTG	CGGAAAATG	CTCTGAACAA	CCTGGACCCA	540
AACACTGAAC	TCAACGTGTC	CCGCTTAGAG	GCTGTGCTCT	CCACTATTTT	TTACCAAGCTC	600
AACAAACGGA	TGCCAACCAT	TCACCAATTC	CATGTGGAGC	AGTCCATCAG	CCTCTCTCCT	660
AACTTCTCTG	TTCAGCGGTT	TGATCCGGAA	GGCCATGGTA	AAATTTCAAT	ATTGCTGTCT	720
AAAATGGCTT	TAGCCACATT	GTGTGGAGGG	AAGATCATGG	ACAAATTAAG	ATATATTTTC	780
TCAATGATTT	CTGACTCCAG	TGGGGTGATG	GTTCATGGAC	GATATGACCA	ATTCCCTCGG	840
GAAGTTCTTA	AACATCCCAT	GGCAGTTTTC	GAAGGTCTTT	CATTGGGTTA	CACAGAACAG	900
TGAGCCAGAT	CCTGTTTCTC	CCAACAGAAA	AAAGTCAAGT	TAAATGGTTT	CTTGGACACG	960
CTTAGTCTAG	ATCTCTCCCT	GCAGTGTCTG	GTCTGGTTGC	CTCTCTGCA	TGCACTAGCA	1020
AATGTGGAJA	ATGTCTTCCA	TCCGTGTGAG	TGTTCTTACT	GCCACAGTGA	GAGTATGATG	1080
GGATTTGCGT	ACCGATGCCA	ACAGTGTTCAC	AATTACCAGC	TCTGTCAAGG	CTGCTTCTGG	1140
AGGGGACATG	CCGTTGGTTC	TCTATAGCAAC	CAGCACCAAA	TGAAGAGATA	CAGTCTATGG	1200
AAATCACCTG	CTAAGAAGCT	GACTAATGCA	TTAAGCAAGT	CCCTGAGCTG	TGCTTCCAGC	1260
CGTGAACCTT	TGCACCTCAT	GTTCCAGAT	CAGCCTGAGA	AGCCACTCAA	CTTGGCTCAC	1320
ATCGTTGATA	CTTGGCCTCC	CAGACCTGTA	ACCAGCATGA	ACGACACCTT	GTTCTCCAC	1380
TCGTTCCTCT	CCTCAGGAAG	TCCTTTTATT	ACCAGGAGCT	CGGACGCTGC	TTTGGGTGGA	1440
TGCGTCTAGA	TGGAATACAT	GACTTCTTCT	ACCCTAAAT	ATTCTCATTA	TACTTTGAGC	1500
TGTTCTGGTT	CCTCAGGAGT	GCAATGTACC	CATTAACCCA	AAATATGATT	ATTTCCTTTT	1560
TTTCCATTTC	TGCAATGTTT	TGCAATGTTT	TCTGTGAACC	ACAGTTGGGT	TGTTAAAGC	1620
TCACATTTCT	TCTGTGACCC	ACAGAGATTG	GCCTACGGTT	TCTGTTTGA	GGGTGCTGTT	1680
CAATAAGCT	GTGTACACTA	AATGTCT				

Seq ID NO: 256 Protein sequence
Protein Accession #: AAC50424.1

45
50

1	11	21	31	41	51	
MLSDSGKRG	TMABERQLFA	EMRAQDLDR	RLSTYRTACK	LRFVQKKCNL	HLVDIWNVIE	60
ALRENALNNL	DNTLELVSR	LEAVLSTIFY	QLNKRMPETH	QIHVEQSISL	LLNFLAALFD	120
PEGHGKISVF	AVKMAIATLC	GGKIMDKLRY	IFSMISDSSG	VNVYGRYDQF	LREVLKLETA	180
VFEGPSFGYT	EQSARSCFPQ	QKKVTLNGFL	DTLMSDPPFQ	CLVWLPLLHR	LANVENVFHP	240
VECSYCHSES	MMGFRYRCQ	CHNYQLCQDC	FWRGHAGGSH	SNQHQMKKEYT	SWKSPAKKLT	300
NALSLSLSLCA	SSREPLPMPF	PDQPEKPLNL	AHIVDTWPFR	PVTSMEDTLF	SBSVPSGSGP	360
ETRSSDGA	GGCV					

Seq ID NO: 257 DNA sequence
Nucleic Acid Accession #: NM_004172.1
Coding sequence: 179..1807

55
60
65
70
75
80

1	11	21	31	41	51	
GCGGATTGTT	GCTCCGTTGT	ACCTGCTGGG	GAATTCACCT	OGTTACTGCT	TGATATCTTC	60
CACCCCTTAC	AAAATCAGAA	AAATTTGTGT	TTCTAATACC	AAAGAGGAGG	TTTGGCTTTC	120
TGTGGGTGAT	TCCCAGACAC	TGAAGTGCAA	AGAAGAGACC	CTCCTAGAAA	AGTAAATAT	180
GACTRAAAGC	AATGGAGAAG	AGCCCAAGAT	GGCGGGCAGG	ATGGAGAGAT	TCCAGCAGG	240
AGTCCGTAAA	CGCACACTTT	TGGCCAGAA	GAAAGTGACG	AACATTACAA	AGGAGGATGT	300
TAAAGGTAC	CTGTTTGGGA	ATGCTTTTGT	GCTGCTCACA	GTCAACGCTG	TCAATGTGGG	360
TACAATCCTT	GGATTTACCC	TCCGACCATA	CAGAATGAGC	TACCGGGAAG	TCAAGTACTT	420
CTCCTTTCTT	GGGGAACCTC	TGATGAGGAT	GTTACAGATG	CTGGTCTTAC	CACCTATCAT	480
CTCCAGTCTT	GTCAACAGAA	TGGCGGCGCT	AGATAGTAA	GCATCAGGGA	AGATGGGAAT	540
GCGAGCTGTA	GTCTATTATA	TGACTACCAT	CATCATTTCT	GTGGTGATTG	GCATATCAT	600
TGTATCATC	ATCCATCCTG	GGAGGGGCAC	AAGGAAAAC	ATGCACAGAG	AAGGCATAAT	660
TGTACAGTG	ACAGCTGCAG	ATGCCCTTCT	GGACTTGATC	AGGAACATGT	TCCCTCCAAA	720
TCTGTAGAA	GCCTGCCTTA	AACAGTTTAA	AACCAACTAT	GAGAAGAGAA	GCCTTAAAGT	780
GCCTATCCAG	GCCAAACGAA	CGCTTGTGGG	TGCTGTGATA	AACAATGTGT	CTGAGGCCAT	840
GGAGACTCTT	ACCGAATCA	CAGAGGAGCT	GGTCCCAATT	CCAGGATCTG	TGAATGGAGT	900
CAATGCCCTG	GGTCTAGTTG	TCTTCTCCAT	GTGCTTCGGT	TTTGTGATTG	GAAACATGAA	960
GGACAGGGG	CAGGCCCTGA	GAGAGTTCTT	TGATTTCTCT	AACGAAGCCA	TCAAGAGACT	1020
GGTAGCAGTA	ATAATGTGTT	ATGCCCCCGT	GGGTATTCTC	TTCCTGATTG	CTGGGAAGAT	1080
TGTGAGATG	GAAGACATGG	GTGTGATTGG	GGGCGAGCTT	GCCATGTACA	CCGTGACTGT	1140
CATTGTTGGC	TTACTCATTC	ACGCAATCAT	CGTCTTGCCA	CTCCTCTACT	TCTTGGTAAC	1200
ACGGAATAAC	CTTGGGTTT	TTATGGAGG	GTGCTGCAAA	GCACCTATCA	CCGCTCTGGG	1260
GACCTCTTCA	AGTTCTGCCA	CCCTACCCAT	CACCTTCAAG	TGCCCTGGAAG	AGAACATAGG	1320
CGTGACAGAG	GATTCTGCTC	CCCGTAGGGA	GCCACCATTA	ACATGGATGG		1380
GACTGCCCTC	TATAGGCTTT	TGGCTGCCAT	TTTCATTGCT	CAGATTACAA	ACTTTGAAC	1440
GAACCTTGGC	CAATATTATA	CAATCAGCAT	CACAGCCACA	GCTGCCAGTA	TTGGGGCAGC	1500
TGGAATTCCT	CAGGCGGGCC	TGGTCACTAT	GGTCAATTGT	CTGACATCTG	TCCGCCCTGCC	1560
CACAGACGAC	ATCAGCTTCA	TCATCGCBGT	GGACTGGTTC	CTGGATCGCC	TCCGGACCCAC	1620

5
10
15
20
25
30
35
40

CACCAACGTA CTGGGAGACT CCCTGGGAGC TGGGATTGTG GAGCACTTGT CACGACATGA 1680
 ACTGAGAAGC AGAGATGTTG AAATGGGTAA CTCAGTGATT GAAGAGAATG AATGAAGAA 1740
 ACCATATCAA CTGATTGACAC AGGACAATGA AACTGAGAAA CCCATCGACA GTGAAACCAA 1800
 GATGTAGACT AACATAAAGA AACACTTTCT TGAGCACCAG GTGTAAATAA CCATTATATA 1860
 ATCTTTCCAT CTCATTACAG CTCATTGCT CCAGCAAGCC CGTCATCTTC CCTTTCCTCC 1920
 CTTCGTATA GACTGSAAAA TAGTCTCCCA AAACACAAGG GAGGATTTTG GGTGGCCAAA 1980
 GTGTACAATT TTCATCCCAC AATTGAAATT TTTAAATCAT TTCATGTTAG TCTTACCBA 2040
 TAAGGTACA AGATCACAAA TAGTGTGAT CAGATCTTAC AAGTTTATGT GGCACACAAT 2100
 TCCTATAAAT GTGATTTTTT TATATAAGTT AAAGAGACAA ATAGTAGGCT AAAAACATTT 2160
 TAAATCAAC TTTTGAATT TAAAAATCTT TCAGAATACA ATTCAATTTT AGTTTCAAAA 2220
 TGTTAACAC TTGAATTACA ACCGGTTATC AGTTGGACAG TAAGATTTTA TCCCTTTCTC 2280
 TTCTGACTGG TATACCTATT TCATTAGTAG CTAGGTGCAC ATATACATCT AGCAGAGCTG 2340
 TGAGGACAGA CAGAAGGCAA AGTTTCCATG TGGCTTGAG CAAGTCCCAT CTCACCTCTA 2400
 GGCTCAGTGG TCCTCATCTA TAAATGAGG GACTTCCCTA GAAGTCITCA TGGTCTCTTC 2460
 CAGCCAGAG ATCCTGTGAT GTCATGAAG CACCTGCCCT CTGTTTCCCT TCAGAACACC 2520
 CTGTACCATC CATGGAGCAC GAGGCCCTCA GAAAGACAC TCAATGGGA GTGAACATTT 2580
 CTAACTAAG ACAGGATGGC TGTGTGTGTT GGTCAACAGG TCCTGTGAGC AAGTGCAGG 2640
 TTATGCAAGT CGCCAGGAG GAGGCCATTC CAGGAGTGGG ATTATTCATC AAACCTTTTG 2700
 CCCAGTTTAT CCATATGGGG GAAGTATTC CTCTTTCTCT ACTCTGGGAA GAATGCTCTC 2760
 TGCCACTCCT CACTGTATGA TAGACTTCGA AAACAGATGA GAAGACTAGC AGCTAGCAAG 2820
 GGTGCTTGTG GTCACACTGT GGAACACTAA AGAGCTAGGA AAGAGTTGAG CACAGGCAAC 2880
 ATTACAAACA AAGGATTGTA AAACACCAAG AGTACAGGTC TTCTTTAAGG AAGAATAAAA 2940
 AAGAAGAGGT TCATTTTTCT GCGTTTCTT TCACTCTGAA ACACCTTTTC TCGAGTCCAA 3000
 AATCATTCOC CCGGTGAAGT CIGCTTACCA AAACATAAGA CAGCTTATAT ATTTGAAAGA 3060
 AGTCAAATGA ATGAGCTCTC TATAGAGT CCATGAGTTG AGTGGGTATT TCTTATTGTA 3120
 AAGTGTTTT CTTAATCAA AAGTCCCTAG AATGAGGAA ACAAAATATT TATTGTGTTT 3180
 GGAATCCCACT TTATCAAAAT ATTCAAACT TTCAGCTGGA GTGGGGTTTG CTTTGTGTTT 3240
 GTTTGTGTC ATAGAGAAA TGGTAGAAGA TGAATCAGTA TGAAGACACT GTCAATGAGG 3300
 TTATGAGAAA AAAACAGCAG GGGCATTAGT TTCAGGCAAG GCAGCTCCCA GGTTTAGAGA 3360
 TTAATTTTAA CCCCCTAAGG AATATCCAGT CAAAGACGCT GAGTGGGAGC TGTGAGGCA 3420
 TAGCAGCTGT GTTGTAGTTT CTGGCTGAAA ATGCTAAGA ATGACTTAA TTATGCTAAC 3480
 AAATGAAAA ATCTAGACAT AGATCCTCTG ATATACAATT AGAGATATT TTATATAGAC 3540
 CCCAGCATT CTGTGCATAA AAGTTAACAT TAGGCTGTGG TGCAGTAAAC ATTATATGTC 3600
 GAGGCTCTAT TTCGGAATA CACTACAAT GTTAAAGTAC GTGGCTGTCC TCTTAAGACA 3660
 CTAGTAGAGC AAAGACTTAA TCAATCAAC TTAATCTGT TACACAATAT GTGTTTTTAA 3720
 ATATACTAAC CATTTCTTAT GGAAGGTCC TGTGGGAGC CCATCCTCTC GCCAGGCCAT 3780
 CACAGGCTCT GCATACACAT GCATCAGTG TGGACTGGGA AGCATTACTT TGTAGATGTA 3840
 TTTTCAATAA AGAAAAAAT AGTTTACAT T

Seq ID NO: 258 Protein sequence
 Protein Accession #: NP_004163.1

45
50
55

1 11 21 31 41 51
 MTKSNGEPEK MGRMERPOO GVRKRTLLAK KKVQNTRED VKSYLFRNAF VLLIVTAVIV 60
 GTILGPTLRP YMSYREVYKY PSFPGELLMR MLQMLVPLI IESLVGMAA IDSKEAGKMG 120
 MRVAVVYMTI TIIAVVIGII IVIIHHPGK TKEMMEREGK IVERTAADAF LDLIRNMFPF 180
 NLVEACFKQF KINYEKRSFK VPIQANETLV GAVINNVSEA METLIRITRE LVPVPGSVNG 240
 VNALGLVVFS MCFPGVIGNM KBQQAALREF FDSINBAIMR LVAVIMWYAP VGLFLFIAGK 300
 IVEMEDMGTI GCGQLMYTUT VIVGLLIHAV IVLFLYFLV TRKNPWFVIG GLLQALITAL 360
 GTSSSSATLP ITFKCLEEMN GVDKEVTRFV LPVGCATINMD GTALYEAALAA IFPIAQVNNE 420
 LNFQIYITIS ITATAASIGA AGIPQAGLVT MVIVLTSVGL PDDITILIIA VDMFLDLRLT 480
 TTNVLGDSLG AGIVEELSRH ELKNRDVEMG NSVIEENEMK KPYQLIAQDN ETEKPIDSET 540
 KM

Seq ID NO: 259 DNA sequence
 Nucleic Acid Accession #: NM_021948.1
 Coding sequence: 48..2783

60
65
70
75
80

1 11 21 31 41 51
 TGTGGCACTG CTGCGTACC CAACCCAGC CCTGGGTAGC CTGCAACATG GCCCAGCTGT 60
 TCCGCCCCCT GCTGGCAGCC CTGGTCTGCG CCCAGGCTCC TGCACTTTTA GCAGATGTTT 120
 TGAAGGAGA CAGCTCAGAG GACCGGCTT TTGCGTGGG CATCGCGGGC GACCGGCCAC 180
 TGCAAGGCGT GCTCGGCGGC GCGCTCACCA TCCCTTGCCA CGTCCACTAC CTGCGGCCAC 240
 CGCCAGCCG CCGGCTGTG CTGGGCTCTC CGCGGTCAA GTGGACTTTC CTGTCCCGGG 300
 GCCGGAGGC AGAGGTGCTG GTGGGCGGG GAGTGGCGT CAAGGTGAAC GAGGCTTACC 360
 GGTTCGCGT GGCATCGCT GCGTACCCAG CGTGCCTCAC CGACGCTCC CTGGGCTGAA 420
 GCGAGCTGCG CCCCAACGAC TCAGGTATCT ATCGCTGTA GGTCCAGCAC GGCATCGATG 480
 ACAGACGCA GCGTGTGGAG GTCAAGGTCA AAGGGTGTCT CTTTCTCTAC CGAGAGGGCT 540
 CTGCCCGCA TGCCTTCTCC TTTTCTGGG CCCAGGAGC CTGTGCCCG ATTGGAGCCC 600
 ACATCGCCAC CCGGAGCAG CTCTATGCG CCTACCTTGG GGGCTATGAG CAATGTGATG 660
 CTGGCTGCT GTCGGATCAG ACCGTGAGGT ATCCCATCCA GACCCACAG GAGGCTCTTT 720
 ACGAGACAT GATGCGCTTC CCGGGGTCC GGAACATGAG TGTGGTGGAC CCGATGACC 780
 TCTATGATG GTACTGTAT GCTGAAGACC TAAATGGAGA ATTGTTCTCT GGTGACCTTC 840
 CAGAGAAGCT GACATTGGAG GAAGCACGG CGTACTGCTA GGAGCGGGT GCAGAGATTG 900
 CCACCAAGG CCAACTGTAT GCAGCCTGG ATGGTGGCTC GACCACTGC AGCCAGGGT 960
 GGTAGCTGA TGGCAGTGTG GGTACCCCA TGTTCACAC CAGCCAGCGC TGTGGTGGGG 1020
 GCTTGCTGG TGTCAAGACT CTCTCTCTCT TCCCAACCA GACTGGCTTC CCAATAAGC 1080
 ACAGCGCTT CAGCTCTAC TGCTTCGAG ACTCGGCCCA GCTTCTGCG ATCCCTGAGG 1140
 CTTCCAAACC AGCTCCAAC CCAGCTCTCT ATGGACTAGA GCTATGCTC ACAGTACAG 1200
 AGACCTTGA GGAATGCG CTGCCTCAG AAGCCACAGA GAGTGAATCC CGTGGGCCA 1260
 TCTACTCAT CCATCATAG GAGGAGGAG GAGTGGAG CTCCACTCCA GAAGACCCAG 1320
 CAGAGGCCCC TAGGACGCTC CTAGAATTG AAACACAATC CATGGTACCG CCCACGGGGT 1380

5 TCTCAGAAGA GGAAGGTAAG GCATTGGAGG AAGAAGAGAA ATATGAAGAT GAAGAAGAGA 1440
 AAGAGGGGGA AGAAGAAGAG GAGGAGGTGG AGGATGAGGC TCTGTGGGCA TGGCCACCG 1500
 AGCTCAGCAG CCGGGGCCCT GAGGCCCTCT TCCCACTBA GCCAGCAGCC CAGGAGGAGT 1560
 CACTCTCCCA GGGCCAGCA AGGGCAGTCC TGCAGGCTGG TGCATCACC CTCTCTGATG 1620
 GAGAGTCAGA AGCTTCCAGG CCTCCAAGGG TCCATGGACC ACCTACTGAG ACTCTGCCCA 1680
 CTCGCCAGGA GAGGAACCTA GCATCCCAT CACCTTCCAC TCTGGTTGAG GCAAGAGAGG 1740
 TGGGGGAGGC AACTGGTGGT CCTGAGCTAT CTGGGTCCTC TCGAGGAGAG AGCGAGGAGA 1800
 CAGGAAGCTC CGAGGGTGCC CCTTCCCTGC TTCCAGCCAC ACGGGCCCTT GAGGGTACCA 1860
 GGGAGCTGGA GGGCCCTCT GAAGATAATT CTGGAAGAAC TGCCCAAGCA GGGACCTCAG 1920
 10 TGCAAGCCCA GCCAGTGCTG CCACTGACA GCGCCAGCG AGGTGGAGTG GCGGTGGTCC 1980
 CCGCATCAGG TGACTGTGTC CCGAGCCCTT GCCACATGG TGGGACATGC TTGGAGGAGG 2040
 AGGAGGGGT CCGCTGCTTA TGTCTGCTG GCTATGGGG GACCTGTGTC GATGTGGGCC 2100
 TCGCTTCTG CAACCCCGGC TGGGACGCTT TCCAGGGCGC CTGCTACAG CACTTTTCCA 2160
 CACGAAGGAG CTGGGAGGAG GCAGAGACCC AGTGCCTGAT GTACGGCGCG CATCTGGCCA 2220
 15 GCATCAGCCA ACCCGAGGA CAGGACTTCA TCAACAACCG GTACCGGAG TACCAGTGGG 2280
 TCGGACTCAA CGACAGGACC ATCGAAGCG ACTTCTTGTG GTCGGATGGC GTCCCTCTG 2340
 TCTATAGGCA CCGTACCTT GGGCAGCCTG ACAGCTACTT CCTGTCTGGA GAGAATCTGC 2400
 TGGTCATGGT GTGGCATGAT CAGGACAAAT GGAGTGACGT GCGCTGCAAC TACCACCTGT 2460
 CCTACACCTG CAAGATGGGG CTGGTGTCTT GTGGGCGGCC ACCGGAGCTG CCGCTGGCTC 2520
 20 AAGTGTTCGG CGCCCAAGG CTGCGCTATG AGGTGGACAC TGTGCTTCC TACCGGTGCC 2580
 GGGAGGACT GGGCCAGCGC AATCTGCCGC TGATCGGATG CCAAGAGAAC GGTCTTGGG 2640
 AGGCCCCCA GATCTCTCTG GTGCCAGAA GACCTGCCCG AGCTCTGCAC CCGAGGAGG 2700
 ACCCAGAGG ACGTCAGGG AGGCTACTGG GACGCTGGAA GCGCTGTTG ATCCCCCTT 2760
 25 CCAGCCCCAT GCCAGTCTCC TAGGGGCGAA GGCTTTGAC ACTGCCCGCC ACAGCACTGC 2820
 CCTGTACACC AATTTTCCC TCACACCTTG CGCTCACCAC AGGAAGTGAC AACATGAC

Seq ID NO: 260 Protein sequence
 Protein Accession #: NP_068767.1

30 1 11 21 31 41 51
 MAQLFLPLLA ALVLAQAPAA LADVLEGDSS EDRAFRVRIA GDAPLQGVLG GALTIPCHVE 60
 YLRPPSPERRA VLGGSPRVKWT FLSRGREAEV LVARGVVRKV NEAYRFRVAL PAYPASLTDV 120
 35 SLALSELSPN DSGIYRCEVQ HGIDDSDAV EVKVKGVVFL YREGSARYAF SPSGAQERACA 180
 RIGAHATPE QLYAAYLGGY EQCDAGWLSQ QTVRYPIQTP REACYGDMDG FPGVRNYGVV 240
 DEDDLVDVYC YAEIDLNGELF LGDPPEKLTL EEARAYCQER GASTATTGQL YAAWDGGGLDH 300
 CSPGWLADGS VRYFIVTPSQ RCGGGLPGVK TLFLFFNQTG FPNKHSRENV YCFRDSAQFS 360
 AIFPASNPAF NPASDGLLEAI VTVTETLEEL QLPQEATESE SRGAIYSIPI MEDGGGGGSS 420
 40 PEDPAEAPRT LLEFETQSMV PPTGFSEEEG KALREERKYE DEEREKEEEE EEEVEDEALM 480
 AWPSELSPG FEASLPTEPA AQEESLSQAP ARAVLQPGAS FLPDGSEAS RPFVRVSGPPT 540
 ETLPTPERN LASPSPSTLV EAREVGEATG GFELSGVVRG ESEETGSSSEG APSLLPATRA 600
 PEGTRELEAF SEDNSGRTAP AGTSVQAQFV LPTDSASRGV VAVVPSAGDC VPSPRNGGT 660
 45 CLEKEGVRC LCLFGVGGDL CDVGLRFPNP GMDAPQACY KHFSTRRSWE EAEQTCRMV 720
 AHLASISTPE EQDFINMYR EYQWIGLNDR TIEGDFLWSD GVPLLYENWN PQQPDSTFLS 780
 GENCVMVWE DQGVWDVPC NYHLSYTCRM GLVSCGPPPE LPLAQVFRP RLRYEVDVTL 840
 RYRCREGLAQ RNLEPLRCQE NGRWEAPQIS CVPRRPARAL HPEDPBEGR GRLLGRWKAL 900
 LIPSSPMFG P

Seq ID NO: 261 DNA sequence
 Nucleic Acid Accession #: NM_004386.1
 Coding sequence: 2..3967

55 1 11 21 31 41 51
 GATGGGGGCC CGGTTTGTCT GGGCCTTGGG CCTTTTGATG CTGCAGATGC TGCTCTTGT 60
 GGCTGGGGAA CAGGGCACAC AGGATATCAC CGATGCCAGC GAAGAGGGGC TCACATGCA 120
 GAAGCTGGGG TCTGGGTTCAG TGCAGGCTGC CTGCGCGAG CTGGTGACCT TGCCCTGCT 180
 CTTTACCTCG CAGCCACGGC CAAGCGCAGC CCGAGATGCC CCTCGGATAA AGTGGACCAA 240
 60 GGTGCGGACT GCGTCGGGCC AGCGACAGGA CTGCGCCATC CTGTGGGCTA AGGACAATGT 300
 CGTAGAGGTG GGCAAAAGCT GGCAGGGAAG AGTGTCACTG CCTTCCTACC CCGCGCGCCG 360
 AGCCACGCC ACCTACTTTC TGGGGCCACT GAGGGCCAGT GACTCTGGGC TGACCGCTG 420
 CCAGTGGTGG AGGGGCACTG AGGATGAGCA GGACCTGGTG CCTTGGAGG TGACAGGTGT 480
 TGTGTTCAC TACCGATCAG CCGGGGACCG CTATGCACAG ACCTTCCGCTG AGGCCAGGA 540
 65 GCGCTGCCGT CTCAGCTCAG CCATCATTGC AGCCCTCGG CATCTACAGG CTGCTTTGA 600
 GGATGGCTTA GACAACTGTG ATGCTGGCTG GCTCTCTGAC CGCACTGTTC GGTATCTTAT 660
 CACCCAGTCC CTTCTGTGTT GCTATGGCGA CCGTAGCAGC CTTCCAGGGG TTGGGAGCTA 720
 TGGGAGGCGC AACTCACAGG AACTCTACGA TGTGTATTGC TTTGCCCGGG AGCTGGGGGG 780
 CGAGGTCTTC TACGTGGGCC CGGCCCGCCG CCTGCACATG GCGCGCGCGC GTGCACAGTG 840
 70 CCGCGGCCAG GGTGCCCGCG TGGCCTCGGT GGGACAGCTG CACTTGGCTT GGCATGAGGG 900
 CCGGACCCAG TCGGACCCGG GCTGGCTGGC CGAAGGCGAG GTGCGCTACC CGATCCAGAC 960
 GCGCGCGCGC CCGTGGCGGG GCCCAGCCCG GGGCGTGGCG ACCGTCTACC GCTTGGCTAA 1020
 CGGACCCGCT TCCCTCTCAC CCGCCGAGCG CTTCAGCACC TACTGCTTCC GAGCTCAAT 1080
 CCCCACGTCA CACATCGGAG ACCTAGAGAC CCTATCCTCT GGGGATGAGG GGGAGATTCT 1140
 75 GTACGACAGG GGGCCCCAG TTAGAGAACT GGGAGCCACC CTGGAGGAGG AAGAGGTGGT 1200
 CACCCCTGAC TTCAGAGAG CTCTGGTGTG CAGTGGGGAA GAAGAAACCC TGATTTTGA 1260
 GAGAGAGCAG GAGTCTCAAC AGACCCCTCAG CCTACCCCTT GGGGACCCCA TGCTGGCTTC 1320
 ATGGCCCGCT GCGGAAAGTG GGCTAAGCAC GGTGGCCCC AGCCCTAGCG ACATGGGGGC 1380
 AGGCACTGCA GCAAGTTTCA ACACGAGAGT GGCCCCAACT GACCCTATGC CTAGGAGAG 1440
 80 GGGGCGCTTC AAGGGTTGA ATGGCGCTA CTTCCAGCAG CAGGAACCGG AGCCGCGGCT 1500
 GCAAGGGGGG ATGAGAGCCA CGGCCAGCC CCGCACCTCA GAGCTGCAG TGAACCAAT 1560
 GGAGCTTCG TTGGCCATGG CAGTCAAGGA GATGTGGGCG AGTGGCCAGA GCCGAGGCC 1620
 CTGGCTCAT CTGACCAATG AGGTGGATAT GCTTGGAGCT GGTCTGCTG GTGGCAAGAG 1680
 CTCCCAAGAG CTTGGCTGT GTGCCCTTAC CATGTGCTCA CCGACATCT CAGGCCACAG 1740
 CAGGCCCTT GTCTTGGAGC TAGAGAAAGC CGAGGGCCCC AGTCCAGGCG CAGCCACCC 1800

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75

```

AGACCTGTTT TGGTCCCCCT TGGAGGOCAC TGTCTCAGCT CCCAGCCCTG CCCCCTGGGA 1860
GGCATTCCCT GTGGCCACCT CCCAGATCT CCCTATGATG GCCATGCTGC GTGGTCCCAA 1920
AGAGTGGATG CTACACACAC CCACCCCAT CTCCACCGAG GCCAATAGAG TTGAGGCACA 1980
TGGTGGAGGC ACCGCCACCG CTCCACCCCT CCCTGCTGCA GAGACCAAGG TBTATTCCCT 2040
GCCTCTCTCT TTGACCCCAA CAGGACAGGG TGGAGAGGCC ATGCCACAAA CACCTGAGTC 2100
CCCCAGGGCA GACTTCAGAG AAACCTGGGA GACCCAGCCT GCTCAGGTCA ACAAGCTGA 2160
GCATCCAGT TCCAGCCAT GGCCTTCTGT AAACAGGAAT GTGGCTGTAG GTTTGTCCC 2220
CACTGAGACT CCACCTAGAC CAACGGGCTT CAGGGGTATC CCGGGTCTG AGTCTGGGGT 2280
CTTCGACACA GCAAGAAAGC CCACCTCTGG CTTCGAGGCC ACTGTAGATG AGGTGACAGG 2340
CCCTCTGCCC TCAGTGTACA GCAAGGGCT GATGCAAGT TCCCCTATCG CCCCCTGGG 2400
GAGCCCTGGA GTCTTCTTGG TACCCAAAGT CACCCCAAAT TTGGAGCCTT GGGTTGCTAC 2460
AGATGAAGGA CCCACTGTGA ATCCCATGGA TTCCACAGTC ACGCCGCCCC CCAGTGAATG 2520
TAGTGAATTT TGGGAACCTG GATCCCAAGT GTTTGAAGAA GCGCAAGCA CCACCTTGAG 2580
CCCTCAGGTG GCCTTGATA CAAGCATTTG GACGCCCTTC ACBACCTTG AGCAGGGGGA 2640
CAAGTTTGA GACACAGAG GTCTACACT GGGCTCTCA AGCTCCAAAC CCCACCCAGA 2700
GCCAGAGGAT CAGGTGGAGA CCCAGGGAAC ATCAGGAGCT TCAGTGCCTC CGCATCAGAG 2760
CAGTCCCTCA GGAACCCGG CTGTTCTCTC TGGGACACCG ACTGCAGCCA GTGTGGGCGA 2820
GTCTGCCTCA GTTCTCTCAG GGTACCGTGG GACCCCTCCA GCACCCCTGT 2880
GCCTGTCACT CTGGCATAG AGGACTTGA ACTGGAGGTC CTGGCAGGGA GCCCGGGTGT 2940
AGAGAGCTTC TGGAGAGAG TGGCAAGTGG AGAGGAGCCA GCGCTGCCAG GGAACCTTAT 3000
GAATGCAGGT GCGGAGGAGG TGCACTCAGA TCCCTGTGAG AACAAACCTT GTCTTCATGG 3060
AGGGACATGT AATGCCATG GCACCATGTA TGGCTGTAGC TGATATCAGG GCTTGGCCGG 3120
GGAGAACGTG TGCAATGAAG TTGATGACTG CCTCTGCAGC CCCTGTGAGA ATGGAGGCAC 3180
CTGTATTGAT GAGGTCAATG GCTTTGTCTG CCTTTGCTTC CCAGCTATG GGGGCGAGCT 3240
TTGTGAGRAA GACACAGAG GCTGTGACCG CGGCTGGCAT AAGTTCCAGG GCCACTGTTA 3300
CCGCTATTTT GGCACCCGGA GGGCATGGGA AGATGCCAGG AAGGACTGCC GCGCGCGCTC 3360
CGGCCACCTG ACCAGCGTCC ACTCACCGGA GGAACACAGC TTCATTAATA GCTTTGGGCA 3420
TGAAACACAG TGAGTGGGCC TGAACGACAG GATCGTGGAG AGAGATTTC AGTGGACGGA 3480
CAACACCGGG CTGCAATTTG AGAAGTGGCG AGAGAACAGC CCGGCAATTT TCTTCGCGGG 3540
TGGCGAGGAC TGTGTGTGTA TGTGCGCGCA TGAAGCGGG CGCTGGAAAG ATGTCCCTGT 3600
CAACTACAACT CTAACCTATG TCTGCAAGAA GGGCAGAGTG CTCTGTGGTC CCCCCTCCGC 3660
AGTGGAGAT GCTCTACTCA TOGGTGCCCG CAAGGCCAAG AACATGTTC ATGCCACTGT 3720
AAGGTACAGG TGCAATGAAG GATTTGCCCA GCACCATGTG GTCCCATTC GATGCCGGAG 3780
CAATGGCAAG TGGGACAGGC CCCAAATGT CTGCACCAA CCAGAGCTT CACATCGGAT 3840
GCGGGGACAC CACCAACACC CCACACAGC CACCAACACA AATCCCGCAA 3900
GGAGCGCAGA AACACACAGA AACACCCAAC GGAGGACTGG GAGAGGAGC AAGGGAAATT 3960
TTGCTGAAGA ACCGAAAAAA AGRAAGCACA ACACCTTTCC CATGCCCTCT CTGGAGCCTT 4020
CGCTTGGGGA GACAGAACCC AGAGAGAAAC AAGAGAGTCC AGAAGTCCCT GAACCCCAA 4080
CTGTTCTCGC AAAAAAATA TTCTTTTGA CAAGGTCTT CTTTTCTTT TTTTACATAC 4140
ACAGATCTCT CTGGCAGGT GGAGCCAGT GTCTGAAAAG TTCAITTCCT TCTGGCTGAA 4200
CTCTGGGAGT GTGTCCAGC TGAGGGAAGC ACAAGTAGCA AAGTCTATTG GTCTGTCTC 4260
TTGTTTGCCA GGTGATTTG AGCAGGCCCT GATGAGGGTG CATGAGTGT TGTTTGCATT 4320
CACATGAAGT AATGCTTTT CACACAGAA ATTACAGACT AGTCAATGT GGTGAAATC 4380
CTAATTCAGG GAGGAAGCCT GGAGCTAGGG TCAATTAGCT TGGGAATAGA AGGCTACACA 4440
GAAGCAGCAC GTTTTGAAC TTGACAAACG CTCTCCCTTT ACCCTGGACT TCAGCCCAAG 4500
TTCCGTCTTT GGTCTTGGTG GATAAACACA CAGTGTGGAG ATCCACGTA CTGCAATTTA 4560
GGGATGTTTT TAGGACAAAC TCCCTCCATG CCTTCAGAGT TAGGAGTGAG AATGATCAA 4620
GCAATATGTA TGGATGAGT GGAGAGTGA TTGCTAACCC TTCCAGGTCT AGTCCAGCGC 4680
TGAGATTGGG TGGTCTGCA TGTGTGATGA ATCTCTTTCA CACAAATAGA CAGAGGATA 4740
TTTAGGGCTA CTGTGAGGT GATTTCTTCC CCTCCATCT CTCAGGGAGA CAAAGAACCT 4800
CCTCTCTGGA CCAAGGAGGT GCTGCCAAGT TTTCTAGGCC AGTGCACATA CCCAGTCTT 4860
AAGCAGACAT TGGTAGTGCC CTTGCCCTGG GTCCCACTCC TGCCCCACCC CACCTTGTCT 4920
CCTGGCCATT GCTTGGTGGT CTAGAACAC TTAAACTTGG AAGTAGTGAC ACCTACCTGC 4980
GGTCAATATG TAGAGAGATG CTCAGTGTGA AAACCTGAAC ACACAAACAC ACACACACAC 5040
ACATTTTCTT CTGTAGATT TTAATTTTTT AAGTGGGAAA GAACCTCACCT TGCCCTTCTCT 5100
CCCCAATATG GCAACCTGTA AAGGTCTCT CCACACAGG GGGCAGGATC CAGTTCCCTC 5160
ATCTCTGGCA GGAAGATTC ACAGCTTTTC CTCCATGTCT GTTACTGTCT TTAGCAGTCT 5220
CGGTAAATAT CTGTGATCA GGGTTAAAAA AGCACCGTGG AAGATGGGCC TCTTCAGGAA 5280
AGAAAAATAA GCAATGAAT GGTCCACCTA GGGGTTCAGT AAAGAAAGAA ATGTGTTAAC 5340
TGAGCCCTGA TCCCTCTTGG GAAGTAATAA TGACCATGTA CAACTAAGAA GTAGACACCA 5400
TGCTAAAGAC TTACATACAA TCTCTTGAA TCTTCTCAAT AGCCCATTGA CTTAGAAACT 5460
GTACTTTTCC CATTTTACAC ACAGTGAAAC TGAGGCTCAG ATATAAAGGA AAGGTACTGG 5520
CTTGAAGTCA CAACACGAC AGGAGTAAGG ATTTGGAATA AGGATTTGGT CCTGTTTTCT 5580
GGACCAATAT CTACTCTGG CTCTGCTTAC ACTTCTCTC CATCACCAA TCCTTACTCC 5640
AAATCCAGAA GTCAGAGCCA ACTCCCATCT TGGTTCTGAC CCAATCTCTG CTCTGACTC 5700
TGGAGAGGAG ATGTAATAT AATGCAACCC TCATACACAT TTAGGAATG GTTAAGAGT 5760
GTAAACTGAA CCCCATTCT TGTCTCAAT CTCTCCCTCT GTAGACATCT ATCTATTAT 5820
GGTTATTATT CAGAAAACCC AGGGATACAG GTTTGCTTTC TTACTTTGAT AACTCTTCT 5880
AGTTTAAAT ATATAATAA ACACATCTTT GGTCTCTAT GTACACAAA AATTTTCTCT 5940
TGTTTGCGGG GGGCTGGGGA TGCASTGTTT TTTGGGGGGT CTGTTTAT GCTCCCTGCC 6000
CTTGAAGCCC TCAGCCGTTT GCGCTGCCCC CACCTCGGCT CCAATGGTGG AGGGGGCTCT 6060
GGTCTTTTCT AAGTGGGCG GTTGTCTTT TGATCTTCC CTTTGGATG TGCGTGTGTG 6120
TCTGCGTGTG CCAATGCGT GGCACGCATA TGAGTGTGTG TGCGTGTGAA CGGCTTTGGG 6180
TCTGCTGCTG TTTGCTGTGA GCTGCAGTGT TCTGTGGGTC TGTGTAATCT GACACTGTGG 6240
ACATTAAATG ACTTCTTGA CATTTAATA AATTTTTTAA CAGTTCAAAA AAAAAAATA 6300
AAAAAATAA

```

Seq ID NO: 262 Protein sequence
Protein Accession #: NP_004377.1

80

1 11 21 31 41 51
| | | | |
MGAPFVWALG LLMQLMLLFV AGEQGTQDIT DASERGLEMQ KLGSGSVQAA LAELVALPCL 60

5	FTLQPRPSAA	RDAPRIKWTX	VRTASGORQD	LEILVAKDNV	VRVAKSWQGR	VSLPSYPRRR	120
	ANATILLLGL	RASDSGLYRC	QVVRGIESDQ	DLVPLEVTGV	VFHYRSARDR	YALTFABEAOE	180
	ACRLSSALIA	APRHLQAARF	DGFNCDAGW	LSDRTVRYPI	TQSRFGCYGD	RSSLFGVRSY	240
	GRNRNQLYD	VYCFARELGG	EVFVVGPARR	LTLAGARAQC	RRQGAALASV	QQLHLAWHEG	300
	LDQCDPGLHA	DGSVRYFIQT	PRRRCGGPAP	GVRTVYRFAN	RTGFPSPAER	FDAYCFRAHH	360
	FTSQHQDLET	PSSGDEGEIL	SABGPFVREL	EFTLEEEEVV	TPDFQEPLVS	SGKEETLILE	420
	EKQESQQLS	FTPGDPMLAS	WPTGEVWLST	VAPSPSDMGA	GTAASSHTEV	APTDFMPRRR	480
	GRFKGLNGRY	FQQKEPFPGL	QGGMEASAPQ	PTSEAAVNQM	EPFLAMAVTE	MLGSGQSRSP	540
10	WADLTNEVDM	PGAGSAGGKS	SPEFWLWPEF	MVPPSISGHS	RAPVLELEKA	SGPSARPATP	600
	DLFWSELEAT	VSAPSPAPWE	AFVATSPDL	PMAMLRGFK	EWMLPHPTPI	STEANRVEAH	660
	GEATATAPPS	PAETKVYSL	PLSLTPTGQG	GEAMPTTPES	PRADFRETGE	TSPAQVNKAE	720
	HSSSSPFWSV	NRNVAVGFEV	TEIATEPTGL	RGIPGSESGV	FDTAESPTSC	LQATVDEVQD	780
	PWPSVYSKGL	DASSPSAPLG	SPGVFLVPKV	TENLEPMVAT	DEGPTVNPMD	STVTPAPSDA	840
15	SGIWEFGSQV	FEAEESTLS	PQVALDTSIV	TPLTLLEQGD	KVGVPAMSTL	GSSSSQPHFE	900
	PEDQVETQGT	VSAFSPAPWE	SPLGKFAVPE	GIPTAAEVGE	SASVSSGEP	VFWDPSPSTLL	960
	PVTLLGIEDE	LEVLAGEFVG	ESFWEEVASG	EEPALEGTFM	NAGAREVHSD	PCENNPCLEHG	1020
	GTCTNANGTM	GCSCDQGFAG	ENCBIDIDDC	LCSPCENGGT	CIDEVNGFVC	LCLPSYGGSG	1080
	CEKDTBGGDR	GWKFGQHCY	RYFAHRRAME	DAEKDCRRRS	GHLSVSHSPE	ESFPLNSFGE	1140
20	ENTWTGLNDR	IVERDFQNTD	NTGLQFENWR	ENQPDNFPAG	GEDCVVMVAH	ESGRWNVPC	1200
	NYNLFPYVCK	GTVLCCGPEFA	VENASLIGAR	KAKNNVHATV	RYQCNEGFAQ	HHVVTIRCRS	1260
	NGKWDPRQIV	CTKPRRSERM	RGHEHHHQH	HQHHHEKSRK	ERRKHKGHPT	EDWEKDEGNF	1320
	C						

25 Seq ID NO: 263 DNA sequence
Nucleic Acid Accession #: E05 sequence
Coding sequence: 7..2085

30	1	11	21	31	41	51	
	GCOCGATGG	CCAGCACAGG	GAGTATCGAG	CTGGAGCACT	TTGAGGAACG	GGACAAAAGG	60
	CCGCGGCCGG	GGTCGCGGAG	AGGGGCCCCC	AGCTCCTCCG	GGGCGACGAG	CAGCTCGGGC	120
	CCCAAGGGGA	ACGGGCTCAT	CCCCAGTCCG	CGGCACAGTG	CCCAGTGCAG	CTTCTACCGC	180
	ACGCGGACCC	TGCGGGCCCT	CAGCTCGGAG	AAGAAGGCCA	AGAAGGCGCG	CTTCTACCGG	240
	AAOCGGGACC	GCTACTCTCA	GGGCTTGGTG	TTTGCCATCT	CCAGCGACCG	CTTCCGGTCC	300
35	TTGATGCGC	TCTCATAGA	GCTCACCCGC	TCCCTGTGCG	ACAACGCGAA	CCTGCCCGAG	360
	GGTGTCCGCA	CTATCTACAC	CATCGACGGC	AGCCGGGAAG	TCACCAAGCT	GGACGAGCTG	420
	CTGGAAGGTG	AGAGTACGTT	GTGTGCATCC	AATGAACCAT	TTCTGTAAGT	CGATTACACC	480
	AAAAATATTA	ATCCAAACTG	GTCTGTGAAC	ATCAAGGGTG	GGACATCCCG	AGCGCTGGCT	540
40	GTCTGCTCCT	CTGTGAAAGG	TGAAGTAAAG	GAAAGTAAAG	ATTTTCATCA	ACCCAGGTTA	600
	GTGACTGTGA	TTCTGAAGTG	AGTGAAGCCT	AGAAAAGCCG	TGCGGATCCT	TCTGAATAAA	660
	AAGACTGTCT	ATTCTCTTGA	ACAAGTCTTA	ACAGATATCA	CCGAAGCCAT	TAAACNAGCC	720
	TCAGGAGTGG	TCAAGAGGCT	CTUCACCCCTG	GATGGAAGC	AGGTGAGAGT	TAGTGTGTG	780
	CATCTGCGAG	ACTTTTITGG	TGATGACGAT	GTITTTTATTG	CATGTGGACC	AGAAAAATTT	840
45	CGTATGCCCC	AGATGACTTT	TGTCTCTGGT	CATAGTGAAT	GTCTGTCTCT	GAAGTCACTC	900
	TATCTGCBAT	CCTCAGCTGT	TAAGTATCT	GGATCCAAAA	GCCTTGGGCG	CTCTCGACGC	960
	AGCCAGATTT	CTGCTCATGG	CAGATCTTCT	TCCAATGTAA	ACGGTGGAGC	TGAGCTTGAC	1020
	CGTTGCATAA	GTCTCTGAAG	TGTGAATGGA	AACAGATGCT	CTGAATCATC	AACTCTTCTT	1080
	GAGAAATACA	AAATTGGAAA	GGTCATTGGT	GATGGCAATT	TTGCAGTAGT	CAAGAGTGT	1140
50	ATAGACAGGT	CCACTGGAAA	GGAGTTTGCC	CTAAAGATTA	TAGACAAAGC	CAATGTGTGT	1200
	GGAAAGGAAC	ACCTGATTTA	GAATGAAGTG	TCAATAGTGC	GCCGAGTGAA	ACATCCCAAT	1260
	ATCATATTGC	TGGTCCAGGA	GATGGAACA	GCAACTGAGC	TCTTTCTGGT	GATGGAATTG	1320
	GTCAAAGGTG	GAGATCTCTT	TGATGCAATT	ACTTGTGTGA	CCAAGTACAC	TGAGAGAGAT	1380
	GGCAGTGCCA	TGGGTGACAA	CTTAGCCCAT	GCCTCAGGT	ATCTCCATGG	CCTCAGCATC	1440
55	GTGCACAGAG	ACATCAAAAC	AGAGATATCT	TTGTGTGTGT	AATATCTTGA	TGGAACCAAG	1500
	TCCTTGAAGC	TGGGAGACTT	TGGGCTTGCG	ACTGTGTTAG	AAGGCCCTTT	ATACACAGCT	1560
	TGTGGCACAC	CCACTTATGT	GGCTCCARAA	ATCATGCTGT	AAACTGGCTA	TGGCCTGAAG	1620
	GTGGACATTT	GGGCACTGG	TGTGATCACA	TACATACITC	TCTGTGGATT	CCACCATTC	1680
	CGAAGTBAGA	ACAACTCTCA	GGAAGATCTC	TTGACCAGA	TCTTGGCTGG	GAAGCTGGAG	1740
60	TTTCCGGCCC	CCTACTGGGA	TAAACATCAG	GACTCTGCCA	AGGAATTAA	CACTCAAAATG	1800
	CTTCAAGTAA	ATGTTGAAGC	TGGGTGTACC	GCGGGACAAA	TCTTGAGTCA	CCCTGGGTGT	1860
	TCAGATGATG	CCTCCAGGGA	GAATAACATG	CAAGCTGAGG	TGACAGGTAA	ACTAAAACAG	1920
	CACCTTAAAT	ATGCGCTCCC	CAACAGAAAC	AGCACTACCA	CCGGGCTCTC	GTCTCATATG	1980
	GTGAGTGGAA	GGCGGCGAGT	CTGGCCTGAC	TGCGGAGCCG	GCCTTGAAGT	TTTTGAATTA	2040
65	GGTAGCCCGG	AGCTGCCCTC	ACATGGAAGT	TGGTGCCCTC	CGTAGTCCCT	TTTCATATGA	2100
	AGATTGGCTT	GGCATGTGGA	GGGCACTCAT	TGGGCAACTC	CCAGGCTTTG	GGCACTGTGT	2160
	GGAGGGGCTT	GTGTAGGGAG	CAGCAGGCTT	GGTGTGAGGG	GTCCAGGCGT	CAAGGAGCTC	2220
	CTGGCTGGGC	CTCTGCGGCA	GCTGCTTCCA	CTCTGTCTCT	TGCTTCTCTA	TCTAGAGAGA	2280
	CTCCCAAGCC	CTGGAGGGGT	GTGTGTGTGT	AGGAATTAAC	TCCCTGCCTA	CCCCAAGGCC	2340
70	TCAGAAATAG	ATTATTAGAG	ATGTGAATTA	TTCTTTGAGA	CTTGGGATAA	GAAACAGCCA	2400
	AAGCTAAACA	TATTTCAAGT	TTAAAAAATC	AGTGTTTTAT	AAAAACACAGT	TGGGGCTTTT	2460
	TAAAGGTACA	TATCAAGGGA	AAAAAATATA	TATTCATTTT	TCAGGGTTGG	TAAACATTTA	2520
	TGAGATGTCA	GTGACAAAGA	TGGCCTTAT	TTTTTCAGCC	TTTTCTTCTT	CCAAATGTTT	2580
	TCTTAAGGCA	ACTCTCTCAA	ATACATAAAC	ACAACAAATT	AAATGAAAA	GTGACATGAG	2640
75	AGTAAATGAA	TCAAAAGGAA	AAAACATGGA	ACCAGAGGTG	AGGGCAGCAC	ACCCGACGCA	2700
	GCTGTCCAGG	TGGAGAGGCA	TGCAACCCCG	GGCGGGAAGG	CCAGCTCACC	GTGAGCAGGT	2760
	AGAAGCCAGC	CAGCCACCCA	GGCAGGGAGC	TTGGTTCTCC	CCACACACTC	CCAGGAGCAG	2820
	GGAACAGGCG	TGGAGTGGCC	TTTCCAGAG	CTGGAGTTGG	CTGCAGCAGG	TTTGAATCAT	2880
	GACCTGCCAA	GGTGTATGGC	GTCTGAGTTT	CACATCTGGG	CCCCCGCTGA	CCCCACTGAG	2940
80	TCCCTGACAG	TGAAGATGGG	CCACCTCCAC	AGCTCCGTCA	CTCGTACTTG	GGACAGGCGT	3000
	CTCATCTCTC	TAAGAGGTTC	TCTTGTGTTT	CTACCCAACT	AGAAAGGAAA	CAGTGGCATA	3060
	TCTCATGAGT	ACATGGTGTG	CTGAAAGCCT	TACCTAGGAA	GAGCAGGGGT	CTAGATAGAA	3120
	GCTATAAGGA	AGCCACGAC	ATAACCCACA	TCCCCACACC	CCCAACATCC	CCCACTCTCC	3180
	CCACACCCCC	CACACCCCCC	ACATCCCCAC	CATAATTACC	CCCACTCTCA	AATATCTCAT	

Seq ID NO: 264 Protein sequence
Protein Accession #: Eos sequence

5	1	11	21	31	41	51	
	MASTRSIELE	HFEERDKRPR	PGSRRGAPSS	SGGSSSSGPK	GNGLIPSPAH	SAHCSFYRTR	60
	TLQALSSEKK	AKKARFYRNG	DRYFKGLVFA	ISSDRFRSPD	ALLIELIRSL	SUNVNLPGQV	120
	RTIYIDGSR	KVTSLEDELLE	GESYVCASNE	PFRKVDYTKN	INPNWSVNIK	GGTSRALAAA	180
10	SSVKSEVKES	KDPIKPKLVT	VIRSGVKPRK	AVRILLNKKT	AHSFEQVLTD	ITEAIKXASG	240
	VVKRLCTLDG	KQVRVTCVHL	PDFFGDDDFV	IACGPEKFRY	AQDDFVLDHS	ECRVLKSSYS	300
	RSSAVKYSGS	KSPGFSRRSQ	ISAHGRSSSN	VNGGFELDRK	ISPEGVNGNR	CSESSLLEK	360
	YKIGKVIQDG	NFAVVKBCID	RSTGKEFALK	IIDKAKCCGK	EHLIENEVSI	LRVRKHPNII	420
	MLVEEMETAT	ELFLVMELVK	GGDLFDATIS	STKYTERDGS	AMVYNLANAL	RYLHGLSIVH	480
15	RDIKPENLLV	CEYEDGTGKS	KLGDFTGLATV	VEGPLYTVCG	TPYVAPKII	AETGYGLKVD	540
	IWAAGVITYI	LLQFFPPFRS	ENMLQEDLFD	QILAGKLEFP	APYWDNITDS	AKELISQMLQ	600
	VNVEARCTAG	QILSHPWVED	DASQENNMQA	EVTGKLKQHF	NNALPKQNST	TGVSVMVVS	660
	GRRQVWPDCG	AGLEVFEFGS	RELPSHGSHC	LP			

Seq ID NO: 265 DNA sequence
Nucleic Acid Accession #: AB020684.1
Coding sequence: 1..1744

25	1	11	21	31	41	51	
	CCCCCTGTC	ATTAATACAT	TAAAGAGATT	CAATCTTTAC	CCTGAGGTAA	TTTTGGCCAG	60
	TGGGTACCGG	ATTTATACCA	AAATAATGGA	CTTGATTGGT	ATTCAAACCA	AGATATGTTG	120
	GACGGTTACC	AGAGAGAGAG	GACTCAGTCC	TATTGAAAGC	TGTGAAGGAT	TGGGAGATCC	180
	TGCTTGCTTT	TATGTTGCTG	TAATTTTAT	TTTAAATGSA	CTAATGATGG	CATTATTCTT	240
30	CATATATGSC	ACATATTGAA	GTGGCAGCCG	ATTAGGAGGC	CTGGTTACAG	TGTTGTGCTT	300
	CTTTTTCAT	CATGGAGAGT	GTACCCGTGT	AATGTGGACA	CCACCTCTCC	GTGAAAGCTT	360
	CTCATATCCA	TTTCTTGTTT	TTGAGATGTT	GCTAGTGACT	CATATTCTCA	GGGCTACAAA	420
	ACITTTATGA	GGAGCTTGA	TTGCACCTGS	CATTTCCTAT	GTATTTTCTA	TGCTTCTCTG	480
	GCAGTTTGCT	CAGTTTGCTG	TTCTTACTCA	GATTGCATCA	TTATTGTCAG	TATATGTTGT	540
	CGGGTACATT	GATATATGTA	AATTACGGAA	GATCATTTAT	ATACACATGA	TTTCTCTTGC	600
35	ACTTGTGTTT	GTTTGTAGTG	TTGGGAACCT	AATGTTATTA	ACTTCTTATT	ATGCTTCTTC	660
	TTTGGTAATT	ATTTGGGGTA	TTCTGGCAAT	GAACCCACAT	TTCTTGAAAA	TAAATGTATC	720
	TGAACCTTAG	TTATGGGGTA	TTCAAGGATG	TTTTTGGTTA	TTTGGAACTG	TCATACTTAA	780
	ATACTTGACA	TCTAAATATT	TGTTATTTGC	AGATGACGCT	CATATTGGCA	ACTTACTAAC	840
40	ATCAAAATTC	TTAGTTTATA	AGGATTTTGA	TACTTTATTG	TATACCTGTG	CAGCGGAGTT	900
	TGACTTTATG	GAAGAGAGAG	CTCCACTGAG	ATACACAAAG	ACATTATTGC	TTCCAGTTGT	960
	TCTGTAGATG	TTTGTGTGTA	TTGTAGAGAA	GATTATTAGT	GATATGTTGG	GTGCTCTAGC	1020
	TAAACACAC	ACACATGTAA	GAAGACACCA	GTTCGATCAT	GGAGAGCTGG	TTTACCATGC	1080
	ATTGCAATTG	TAGCATATTA	CAGCCCTTGG	TATTTTAATT	ATGAGACTAA	AACCTCTCTT	1140
45	GACACACAC	ATGTTGTGTA	TGGCATCACT	GATCTGCTCA	AGACAGCTAT	TTGGATGGCT	1200
	CTTTTGCAAA	GTACATCTCG	GTGCTATTGT	GTTTGTCTTA	TAGCAGCAAA	TGTCAATACA	1260
	AGGTTGAGCA	AATCTGCAAA	CCCAGTGGAA	TATTGTAGGG	GAGTTGAGCA	ATTTGCCCCA	1320
	AGAAGAAGCT	ATAGAATGGA	TCAAAATATG	TACTAAACCA	GATGCAGTGT	TTGGGGGTGC	1380
	CATGCCACAG	ATGGCAAGTG	TTAAGCTCTC	TGCACCTTGG	CCCATTGTGA	ATCATCCACA	1440
50	TTATGAAGAC	GCAGGCTTGA	GAGCCAGAAC	AAAAATAGTA	TACTCAATGT	ATAGTCGGAA	1500
	AGCAGCGGAA	GAAGAGAGAG	GAGAACTGAT	AAAGTTAAAA	GTGAAGTATT	ACATTTCTAGA	1560
	AGAGTCATGG	TGTGTAGAGAA	GATCCAGGCC	TGGTTGCAGT	ATGCTTGAAA	TTTGGGATGT	1620
	AGAAGATCTC	GCCATGCTGC	GGAAACTCTC	CTTATGTATC	CTCTTGCTGA	AGGATTCCAA	1680
	ACCTCAGCTT	ACCACGTGAT	TCCAGAACAG	TGTTTACAAA	GTCCCTAGAG	TTGTAAAGAA	1740
55	ATGACTGCTA	CATGACCTGC	TGCCCTAGGA	GAATACATCT	TGTAATGGTT	TTAATGTTTT	1800
	GCTAAGTCAT	GTGTGTGTTA	TATCCCAAAA	ACTTTTATAG	GTAACTGTTT	TCAAAATAGAA	1860
	AACGTTTTAT	TGTTCTCAAT	TGAATGTCTA	TCTAATTATA	AAAATGACTT	ACACCTTTAT	1920
	CAATTGGTTA	CTAATTTCAAT	GCACCCCTTA	AAATTTGCTA	TGCAATGAG	TATATGCTTG	1980
	TACTTGACTT	TATATTTTGT	GCTAAAGTGA	GCAAGCTAC	CTGTATAAAG	AAAACACAGT	2040
60	GGGTTGTGAC	AAGGATGACA	TGAAATATCA	GGACAATCTT	GACAATGTAG	GGGCTGATTT	2100
	TATAGTGTAA	GAACTATTA	TGCCCTTGGC	TTCTTTTTC	TGCCCTCTGG	TCTTGTCTTT	2160
	TGGACATTTT	AGTGATTTGA	AGTTCTTGGG	TCAATGTGAG	CCCTGTCTAT	AACCTGAGTT	2220
	ACAGTAGATG	GGGAGACAT	GGAGGTTTGG	CTATATAAAA	CTATCTGTTT	GTTTACTTTC	2280
	CTTGTGGGCT	TTTGTCTCTC	TGTTCTCTTG	TAAATGAAGC	TTTTCTCTGC	CATTATTAAAT	2340
65	CCAACTCTTT	GGACCTTGGT	GTTAGGAAAT	TCCCTTAACT	TCCAGGCATA	TGGCATTATC	2400
	GTGTCTCTTT	CTCTCTCTCT	CTTCTCTCT	CTCTCTCTCT	CTTCTCTCTCT	TTTTCTCTCT	2460
	AATAAGTACT	GTTTACTCAT	TGATTTGCTT	ATCAAGTACT	TATCTCTGTT	TTTAAAAAAA	2520
	ATTAATGGTA	ACTGTATTTT	TCTCATTTT	AGCATTATTC	AAATGTTTAT	ATTTTAATAC	2580
	CTTTAAACCA	CTTTAAAGTT	TTTTCATGTT	TAATTTATGT	TTTAAAGAAA	ACTATTTTGA	2640
70	ACAACCCCAA	ATATAGTGCA	TCTAGAAACT	AATGTATATT	TGATTAGACA	TCAATTATAG	2700
	TGGAAACGTA	GACTGTAGTA	CATGGTAAAT	TTTCTTTTAC	TATTAAGATA	CAATAAAACA	2760
	TGACTAATTT	TGCTGTCAAA	AATGTAAAGA	ATATGTATAA	ATGGAGTTTT	TTATATTTTA	2820
	CTTTAAGAT	TGCCCTGCTT	TAAATAGACA	AAGCCTTAAG	CCCTATGTGA	TAATTTTGGT	2880
	TCTAAAAACC	ATCATTTTCA	TATAAGGAAT	AGATATATTT	OGTCTCTCTC	TTTAGTTTTT	2940
75	TTCTTCTCAT	TTATTTTTAT	TTTGAAGAAAT	TTCTACACCT	TCCTTGAATT	CCTTGTATGA	3000
	ATTTTGTGTT	CTTAAAGATT	AATTTGTGTT	AAATGAGATT	CTTCAAAAGC	ATGAAACCTC	3060
	ATAGCTCTGA	GAAAGAGTTT	TAGGGTTTAA	AATCTAAGC	AAAGCTGTAC	TATGGCTGAC	3120
	AGACTACACA	TTTAATTTAA	CAGCTTCTCT	TTCTTAAACA	CAGGCAGATT	AACCTCATAG	3180
	TGGATTGTCC	TTCAAGCCTT	AGTCTCTCAG	CATGTTTCTT	GGTCCCACTT	CCTGGAGGCC	3240
80	GCTGTTCCCT	TTCTACCTTC	TTACCAAGAG	CCAAGGGCAG	GCTTGTGCCC	GGGGAAGCCG	3300
	CAGCTTGCCT	ACATAAGTCA	GCTGCAAAAG	CAGGAGGATG	TGCCCTCAGA	GAAGCACCCG	3360
	CCCCAGTCT	TGTGCCAGCG	CCTAGAGCCG	CAGCTCCAGC	GGAATGCTCT	TCCCTGGAGG	3420
	CAGCCACAGA	GAGGACTCTT	GGCAGGCTTC	TTCAAGTTTG	TGGCCACTGT	TTCTCATTTG	3480
	CTGGTTGACT	GTTTTTATTT	CTTAGGCTTT	TGCTAGTTTT	AGAAAATAGG	GAAGCAGCCC	3540
	TGATTTTGTG	GATTAAAGAG	AACATTGAG	CGATGATGCA	CAACAGTCCA	GGAAAATGGG	3600

CGGTGGACAC TTGAGGCTGA GGATGGGAGT TGACATGAGC AGGGAGAGGG AGGTGCGCGC 3660
 TGCTTATCTG TGATTGTTGC TCACCTGAGT GTGGCTGATT GTGTACATCC AGCAGTTACA 3720
 ATTTTAAAAA ATTATACCTT TACATTATTT TTATATTTTT CTCACCCCCA GTAATTTCCT 3780
 TCCAAAGAGG TTCAATAGTA ATAAGTAGAA ATTCGTGATA GGAAGAAAGC ATTAAAAATA 3840
 CTATTATAAC TGCTTCATTG GCTGGGAACC ATTAAAGTA ATATAAATTA GCTTTTTCCT 3900
 GAAGGATCCT TTTGTAGCAG TGTATTAGAA TGTAACCCCC AGCAAAATAT GGCTATATAT 3960
 TAGGGGAGCC AGTTTGGAGC AGAGGCTGGA AGGTCCCTGC TATGCAGCCG TGGCCACAGC 4020
 TCGCAGCCCA AGCACTGTGG AGCATCCACA CCTTTGATGG CAATGCAGAT TGGTAGCAGG 4080
 TTCCATAGGC GTACAAAACA GTATTAAAGC TCAGTGTGTT GCATATGTTT AGCATTTACA 4140
 AATATTTTTG CTTTAGTATG AGGAAAGTAA GGATGGGCAA AGAAGCGATC AAAATAGCTA 4200
 TTGCTACAAAC ATTTTCGAAA ACAAAATTGG GGCTGTATTT CTTTAAAAAG ATAAGCCTCT 4260
 AAAAATGCTT GGCAGAAAAA ATATAGTGTT AAAATAGGCC AGTGATATTA ATGAGAAAT 4320
 GAAAGTATGT ATCAGGAATA AAGTGATATT GCATAGGAGT ATTGTATTTT TATGAATTTT 4380
 ATGCCAGTTG TTTACATGTA CTATATATGT TAAATTAAAA AAAATCATGA GAAATG

Seq ID NO: 266 Protein sequence
 Protein Accession #: BAA74900.1

1 11 21 31 41 51
 PLVINTLKRF NLYPEVILAS WYRIYTKIMD LIGIQTKICM TVTRGEGLSP IESCEGLGDP 60
 ACFYVAVIFI LNLGLMALFP IYGTLYSGSR LGGLVTULCF FPNHGECTRV MWTPLRESF 120
 SYPFVLVQLM LVTHILRATK LYRGSLLALC ISNVFFMLPW QFAQFVLLTQ IASLFAVYVY 180
 GYIDICKLRK IYXHMISLA LCFVLMFGNS MLTSTYASS LWIINGILAM KPHFLKINVS 240
 ELSLWVIGGC FNLFGTVILK YLTSKIFGIA DDAHIGNLLT SKFFSYKDED TLLYTCABEF 300
 DFMKEKTLPLR YTKILLPLVV LVVFAIVVRK IISDMWGLA KQQTHVRKHQ FDEGELVYHA 360
 LQLLAYTALG ILLMRLLKFL TPHMCMVASL ICSRQLFGWL FCKVHDPGAIV FAILAAMSIT 420
 GSANLQTQWN IVGEFENLPQ BELIBWIKYS TKPDVAFVGA MPIMASVKLS ALRPVNVNHF 480
 YEDAGLRART KIVYSMYSRK AAEFVKRELI KLVNMYIILE ESWCVRRSKP GCSEPRINDV 540
 EDPANAGKTP LCNLLVKDSK PHEFTVFQNS VYKLEVVVKE

Seq ID NO: 267 DNA sequence
 Nucleic Acid Accession #: U26744.1
 Coding sequence: 59..1600

1 11 21 31 41 51
 CTTCAAAGAA TATAGATGGT TTTGAAAAGT TCGTCTGTC CCTTCATTGA ATTTTAGAAT 60
 GATTGAAGAT AGTGGGAAAA GAGGAAATAC CATGGCAGAA AGAAGACAGC TGTTTGAGAA 120
 GATGAGGCTC CAGATCTGG ATCGATCCG ACTCTCCACC TACAGAACAG CATGCCAAGCT 180
 TAGGTTTGT CAGAGAAAT GCAATTGCA CCTGGTGAGC ATATGGAATG TCATAGAAGC 240
 ATTGGCGGAA LLMRLLKFL ACAACCTGGA CCAAAACACT GAACCTCAAG TGTCCCGCTT 300
 AGAGGCTGTG CTCGCCACTA TTTTAACTCA GCTCAACAAA CGGATGCCAA CCATCCACCA 360
 AATCCATGTG GAGCAGTCCA TCAGCCCTCT CCTTAACTTC CTGCTTGACG CGTTTGATCC 420
 GGAAGGCCAT GGTAAAATTT CAGTATTGCG TGTCAAAATG GCTTTAGCCA CATTTGTGTG 480
 AGGGAAGATC ATGCAAAAT TAAGATATAT TTTCTCAATG ATTTCTGACT CCAGTGGGGT 540
 GATGTTTAT GGACGTCTAT ACCAATCTCT TCGGGAAGTT CTCAAACTAC CCACGGAAGT 600
 TTTGGAGGTT CCTTCATTG GTTACACAGA ACAGTCAGCC AGATCTCTGT TCTCCCAACA 660
 GAAAAAAGTC ACCTTAAATG GTTCTTGGG CAGCCTTATG TCAGATCTCT CCCCAGAGT 720
 TCTGTCTGTG TTGCTCTTC TGCACTGACT AGCAAAATG GAAAATGTCT TCCATCCGGT 780
 TGAGTGTGCC TACTGCCACA GTGAGAGTAT GATGGGATTT CGCTACCGAT GCCAACAGT 840
 TCACATATAC CAGCTCTGTC AGGACTGCTT CTGAGGGGGA CATGCCGATG GTTCTCATAG 900
 CAACACGAC CAAATGAAAG AGTACACGTC ATGGAATCA CCTGCTAAGA AGCTGACTAA 960
 TGCAATAGC AAGTCCCTGA GCTGTGCTTC CAGCCGTGAA CCTTTGCACC CCATGTTCCC 1020
 AGATCAGCCT GAGAGCCGAC TCAACTTGGC TCACATCGTT GATACTTGGC TCCCCAGACC 1080
 TGTAAACGAC ATGAAACGAC CCGTGTCTTC CCACTCTGTT CCGCTCTCAG GAAGTCTCTT 1140
 TATTACAGG AGCATGCTG AGAGTTCAAA CCGCTTGAT GAAGAACACA GGCTAATTGC 1200
 CAGGTATCGG GCAAGGCTGG CAGCAGATC CTCCTGCTCT CAGCCACCTC AGCAGAGAG 1260
 TGCTCTGAC ATCTCTTCA CCAATCGATG GAATAAGCAG CAAAGGCAGC TGATTGCTGA 1320
 GCTAGAAAC AAGAACAGAG AAATCTTACA GGAGATCCAG AGACTTCGGC TAGAGCATGA 1380
 ACAAGCTTCT CAGCCACGCG CAGAGAGGCG ACAGCAAAAC CCAACCTTBC TGGCAGAACT 1440
 CCGGCTCCCT AGACAGCGCA AAGATGAGCT GGAACAGAGA ATGCTCTGCTC TCCAGGAGAG 1500
 CCGGAGAGAG CTAATGGTCC AGTTGGAGGG TCTCATGAAG CTACTAAAGG AAGAAGAACT 1560
 GAAGCAGGGA GTAAGTTATG TCCCTTACTG CAGGTCTTAA CTAACAGTGG AGGGGCTGC 1620
 CGACCTGCGG TTTTCTCAT GCTTTGCTC TAATGTATGT TCATGCTTCA GTTTGGAAAG 1680
 AGAAAAAGT CATACTAAT TGCTCTTCTT TCAATGTAGT GCTTGAATTG AGATATATAA 1740
 ATTTAGCAT TTTTATAACT ATCACTACTA TCACATCRA AAGAGAACT ATGACATCTT 1800
 TTAGAAAGG GAACGAATG TCATTTATG GAACATTTT AGATCCCGAG AGGTATAAGT 1860
 TTCAAACCA TCTTAGCTTT TCAAGTTGTT GATCAGACCC TTCTCTTAAC AGAGAGATAC 1920
 CACAGTCAC AGAGATACCC TGAGGTTTCT GTCATCCCAA AACCCACAGC ACTCAGAGC 1980
 TAACCTCTAC ACCCACTCAC ACTGTGAGTA TTCAGTTCGG TTTCAATTTA CTGAAAACCT 2040
 GTGAACCTCT TTTTATATA AATCAGGCAA TTAATCCCT TTTCTACCA CAATTAITGA 2100
 GCCTTGTTC CCATGGCTCA CCAAAATGTG CTCATTTTG TGAGAGAAAG ACTGTACTCC 2160
 ATAACTGACT ATTCAGCTCC CATCTTTTG GCTCTTCCC AAAGCAGAA CTCTACTGTT 2220
 GGTGTGACAT AATCTCTTT TAAAAAGTAA CTCTCAGCT TTTCTTAGC ACCAGAGCCT 2280
 TTCGGCTCCG GAGACGAGA GGGTCATTAC ATACTTTTT TTTTCTCTG AAATAGGGGC 2340
 ATGTGTACTT TATAGCTTAA ACTGGAGCTG TCTGAACCTG TGGTCAGGCT CAAGAGCCAG 2400
 CAGGGGAGC AGCAACTC

Seq ID NO: 268 Protein sequence
 Protein Accession #: AAC50426.1

1 11 21 31 41 51

MIEDSGKRGN TMAERRQLFA EMRAQDLDR I RLSTYRTACK LRFVQKKCNL HLVDIWNVIE 60
 ALREXALNNL DPNTIELNVSR LEAVLSTIFY QLNKRMPTTH QIHVEQSISL LLNELLAAFD 120
 PEGHGKISVF AVKMALATLC GSKIMDKLRY IFSMISDSBG VMVYGRYDQF LRFVLKLPTE 180
 VLEGPSFSGYT EQSARSCEFQ QKQVTLNGFL DTLMSDFPPQ CLVWLPPLER LANVENVPHF 240
 VBCSYCHSES MMGFYRCQQ CHNYQLCQDC FWRGHAGGSH SNQHOMKEYT SWKSEFAKLT 300
 NALSLSLSCA SSREPLEHMF PDQPEKPLAL AHIVDTWPPR FVTSMDITLF SHSVFSSGSP 360
 FITRSMLESS NRLDESHRLI ARYAARLAAE SSSSQPPQQR SAPDISFTID ANKQQRQLIA 420
 ELENKNREIL QEIQLRLLEH EQASQPTPEK AQQNPTLLAE LRLLRQRKDE LEQRMSALQE 480
 SRRELMVQLE GLMKLLKEBE LKQGVSYVFEY CRS

Seq ID NO: 269 DNA sequence
 Nucleic Acid Accession #: NM_001276.1
 Coding sequence: 127..1278

1 11 21 31 41 51
 AGTGGAGTGG GACAGGTATA TAAAGGAAGT ACAGGGCCTG GGGAGAGGCG CCTGTCTAGG 60
 TAGCTGGCAC CAGGAGCGGT GGGCAGGGA AGAGGCCACA CCTGCCCTTG CTCTGCTGCA 120
 GCCAGAAATGG GTGTGAAGGC GTCTCAAACA GGCTTTGTGG TCTTGGTGCT GCTCCAGTGC 180
 TGCCTGCAT ACAAACTGGT CTGCTACTAC ACCAGCTGGT OCCAGTACCG GGAAGGCGAT 240
 GGGAGCTGCT TCCAGATGCG CTTGACCGC TTCCTCTGTA CCCACATCAT CTACAGCTTT 300
 GCCAATATAA GCAACGATCA CATCGACACC TGGGAGTGGG ATGATGTGAC GCTCTACGGC 360
 ATGCTCAACA CACTCAAGAA CAGGAACCCC AACCTGAGA CTCTCTGTCT TGTCCGAGGA 420
 TGGAACTTTG GGTCTCAAGG ATTTTCCAAG ATAGCCTCCA ACACCCAGAG TGGCCGGACT 480
 TTCATCAAGT CAGTACCGCC ATTCTGCGC ACCCATGGCT TTGATGGGCT GGACCTTGCC 540
 TGGCTCTACC CTGACCGGAG AGACAAACAG CATTTTACCA CCTAATCAA GGAATGAAG 600
 GCGAATTTA TAAAGGAAGC CCAGCCAGGG AAAAGCAGC TCTGCTCAG CGCAGCACTG 660
 TCTGCGGGGA AGGTCAACCAT TGACAGCAGC TATGACATTG CCAAGATATC CCAACACCTG 720
 GATTTTCATTA GCTCATGAGC CTACGATTTT CATGGAGCCT GGGCTGGGAC CACAGGCCAT 780
 CACAGTCCCG TGTTCGAGG TCAGGAGGAT GCAAGTCTCG ACAGATTGAG CAACACTGAC 840
 TATGCTGTGG GGTACATGTT GAGGCTGGGG GCTCTGCGCA GTAAGCTGGT GATGGGCATC 900
 CCCACCTTCG GGAGGAGCTT CACTCTGGCT TCTCTGAGA CIGGTGTGG AGCCCCAATC 960
 TCAGGACCGG GAATTCAGG CCGGTTCAAC AAGGAGGCG GGCCTTTC CTACTATGAG 1020
 ATCTGTGACT TCCTCCCGCG AGCCACAGTC CATAGAACC TCAGCCAGCA GGTCCCTTAT 1080
 GCCACCAAGG GCAACAGGTG GTTAGGATAC GACGACAGG AAGCCCTCAA AAGCAAGGTG 1140
 CAGTACCTGA AGGATAGGCA GCTGGCAGGC GCCATGGTAT GGGCCCTGGA CTTGGATGAC 1200
 TTCCAGGGCT CCTTCTCGCG CCAGGATCTG CGCTTCCCTC TCACCAATGC CATCAAGGAT 1260
 GCATCGCTG CAACGTAGCC CTCGTCTCTG CACACAGCAC GGGGGCCAGG GATGCCCGGT 1320
 CCCCCTCTCG CTCCAGCTGG CCGGAGGCGT GATCACCTGC CCGTCTGAGT CCCAGGCTGA 1380
 GCTCAGTCT CCTTCCCTTG GGGCCTATGC AGAGGTCCAC AACACACAGA TTTGAGCTCA 1440
 GCCTGGTGG GCAGAGAGGT AGGGATGGGG CTGTGGGGAT AGTGAGGCAT CGCAATGTAA 1500
 GACTCGGGAT TAGTACACAC TTGTTGATGA TTAATGGAAA TGTTTACAGA TCCCAAGCC 1560
 TGGCAAGGGA ATTCTTCTCA CTCCTTGCCC CCTAGCCCTC CTTATCAAAG GACACCATTT 1620
 TGGCAAGCTC TATCACCAAG GAGCCAAACA TCTTACAAGA CACAGTGACC ATACTAATTA 1680
 TACCCCTTCG AAAGCAGCTG TGAACCTTC ACTTAGGAAC GTAACTCGTG CCCCTATCCT 1740
 ACTTCCCTCT CCTAATCTCA CAGCTGCTCA ATAAAGTACA AGAGTTTAA AGTGTGTGG 1800
 CGCTTTCGCT TGGTCTATCT TTGAGCGCCC ACTAGACCCA CTGGACTCAC CCCCCCATC 1860
 TCTTCTGGGT TCTTCTCTCT GAGCCTTGGG ACCCCTGAGC TTGCAGAGAT GAAGGCCGCC 1920
 ATGTT

Seq ID NO: 270 Protein sequence
 Protein Accession #: NP_001267.1

1 11 21 31 41 51
 MGVKASQTGF VVLVLQCCS AYKLVCTYS NSQYREGDGS CFPDALDRFL CTHIISYFAN 60
 ISNDHIDTWE WNDVITYGML NTLKKNENML KILLSVGGWN FGSQRFKIA SNTQSRRTFI 120
 KSVPPFLRTH GFDGLDLAWL YPGRHDKQHF TLLIKEMKAE FIKEAQPGKK QLLLSAALSA 180
 GKVTIDSSYD IAKISQHLDF ISIMTYDFRG ANRGTTGHS PLFRGQEDAS PDRFENITDA 240
 VGVMLRLGAP ASKLVMGIPT FGSRFLASS RTGVGAPISG PGIPGRFTKE AGTLAYVEIC 300
 DFLRGATVHR TLGQVPEYAT KGNQNVGYDD QESVKSRVQY LKURQLAGAM VWALDLDDFQ 360
 GSPCGQDLRF PLTNAIKDAL AAT

Seq ID NO: 271 DNA sequence
 Nucleic Acid Accession #: NM_006474.1
 Coding sequence: 181..669

1 11 21 31 41 51
 GCTGCTAGG GTCTGGAAG CTCGGGCACC CTCCCTCTCC GGGGCTCCTG CTCCACCCOC 60
 TCGGSCCCOC CCACCGTCCG GCTCCTCCAG GCTGGGCTCG TGGCCGCGGT GCTTTTAATT 120
 TCCCCCCAGC TCAGAACTCT GCTGCTCGGC CCCAGGAGA GACAAACTC AACGGGAACG 180
 ATGTGGAGG TGTCACTCT CTCTTCTGTT TTGGGAAGCG GTCGCTCTG GGTCTTGCA 240
 GAAGGAGCCA GCACAGGCCA GCCAGAGAT GACACTGAGA CTACAGTCTT GGAAGCGGCG 300
 GTGGCCTAGC CAGGTGCGGA AGATGATGTG GTGACTCCAG GAACCAAGCA AGACCGCTAT 360
 AGTCTGCTCT TGACAACTCT GGTGGCAACA AGTGTCAACA GTGTAAACAG CATTCGCTATC 420
 GAGGATCTGC CAACTCTAGA AAGCACAGTC CACGCGCAAG AACAAAGTCC AAGCGCCACA 480
 GCTTCAAACG TGGCCACCAG TCACTCCACG GAGAAAGTGG ATGGAGACAC ACAGACACAC 540
 GTTGAGAAAG ATGGTTTGTG AACAGTGACC CTGGTTGGAA TCATAGTTGG GGTCTTACTA 600
 GCCATCGGTT TCATTGGTGG AATCATCGTT GTGGTTATGC GAAAAATGTC GGGAGGGTAC 660
 TCGCCTTAAA GAGCTGAAGG GTTACGCGCT GCTTGCCAAC GTGCTTTAAA AAAAGACCGT 720
 TCTGACTCT GTGCGCTGT CCTGAGCTC GTGGGAGAAA GATGACCTG GGAACATTG 780
 CGGGCCCATC CAGATTCCAC GGTGACTTTC CGTTTGCCAA ATTAACCGAG GAAGACCTT 840

TCACCAGATT TGGTTCCTAA ACTTT

Seq ID NO: 272 Protein sequence
Protein Accession #: NP_006465.1

1	11	21	31	41	51	
MNKVSALLFV	LGSASLWVLA	EGASTGQPED	DTETTGLEGG	VAMPGAEDDV	VTPTSEBDY	60
K9GLTLVAT	SVNSVTGIRI	EDLPTSESTV	HAQEQSPSAT	ASNVTASHST	EKVDGDTQTT	120
VEKDLSTVT	LVGIIVGVLL	AIGFIGGIIV	VVMRMMSGRY	SP		

Seq ID NO: 273 DNA sequence
Nucleic Acid Accession #: CAT cluster

1	11	21	31	41	51	
GCGGCCGCCA	GCTTGCAGAG	CCGAAGTCTG	GCCGCGCTCT	TCGACTGCT	GCGCCACGTC	60
CCCGGGGGTG	CCGAGCCCGC	GGGGGGTGAG	GTGGCTGCGC	CGGCGGCCCG	GCTAGGAGGT	120
CGGGGCACTG	GGGGCGCGGG	AGGGGACGTG	GCAGGCCCGC	CGGGGGCCAC	GGCGATCCCA	180
GGGCCACAGA	AGGTCCCGCT	GCGGGCAAGC	AATCTGCCTC	CGTCTTCTTT	CACGGAGCCG	240
TCGGGGGCG	GCGGCGCGGG	GTGTGGCCCG	TCGGGGGCGG	ACGTGAGCTT	GGGCGACCTG	300
GAGAAAGGGC	CGGAGGCCGT	CGAGTTCTTT	GAGCTGCTGG	GGCCCGACTA	CGGCGCCGGC	360
ACGGAGGCGG	CAGTCTTGCT	TGCCGCGGAG	CTCTCGACG	TGTTCCCGCG	CGGAGCCTCC	420
GTACTGCGGG	GACCCCGCGA	GCTGGAGCCC	GGCCTCTTTG	AGCCGCGCGC	GGCAGTGGTG	480
GGAAACCTAC	TGTACCCCGA	GCCCTGGAGC	GTCCCGGGCT	GCTCCCGGAC	CAAAAAGAGC	540
CCCTTGACTG	CCCTCGCGGG	CGGGTGGACC	TTGAACGAGC	CCTTGGAGCC	CCTGTACCCC	600
CGCGCTGCGA	ATTTCCTCCG	GCGGGGAGGA	CGGGCGGGGC	CATTGGCTTT	CTTTGCGCCC	660
CTCTTTCCA	GACTGCGCTT	TGC				

Seq ID NO: 274 DNA sequence
Nucleic Acid Accession #: Bos sequence

1	11	21	31	41	51	
CAAGAGAGGC	GGGCTCCAGC	TCGGGGGGTC	CCCGCAGTAC	GGAGGCTCCG	GCGGGGAACA	60
CGTCGAGAGG	CTCGGCCGCA	AGCAAGACTG	CGGCTCCGCT	GCGGGCGCGG	TAGTCGGGCG	120
CCAGCAGCTC	AAAGAACTCC	ACGGCCCTCG	CGCCCTTCTC	CAGGTGAGCC	AAGCTCAGGT	180
CGGGCCCGGA	CGGGCCACAC	CCGCCGCGGC	CTGCCCGGGA	CGGCTCCGTC	AAGAAGGACG	240
GAGGCAGATT	CGGTGCGCGC	AGCGGGAGCT	TCCCTGGCCC	TGGATCCGCT	GTGGCCCGCG	300
CGGGGCTGCG	CAGTCCCGCT	CCCGGCGCCC	CAGTGGCGGC	ACCTCCCTAG	CCGGCCCGCG	360
GCGCAGCCAC	CTCACCCTCC	GCGGCTCGG	CACCCCGGGG	GAGTGGCGCG	AGCGAGTGA	420
AGAGCGCGGC	CAGACTTCGG	CTTTCGAGC	TGGCGGCGGC			

Seq ID NO: 275 DNA sequence
Nucleic Acid Accession #: NM_001118.1
Coding sequence: 74..1651

1	11	21	31	41	51	
AGCCCAAGAG	CACATTGGGG	CTGACCTGCC	GCTGCTGTCA	GTGGGAGGCG	AGTGGTGTCT	60
GCACAGAGAT	GTCTAGGCTG	GTGTGCTGCA	CGTTTCCCTG	GCTGCTCACT	GCGGGGCGCT	120
TCGCTGGGGC	CGGGGCAAGC	TCCGCAAGAG	ACGCGCAGCT	TGCAAGTCCG	CGGCCAGAG	180
ACACATTGGG	GCTGACCTGC	CGCTGCTGTC	AGTGGGAGGC	CAGTGGTCTT	GGCCAAGAG	240
TGTCTAGGCT	GGTGTGCTGC	ACGTTTCCCT	GGCTGCTCTC	CTCTGCTGCT	CTATGGCCCG	300
TGCCATGGAT	CTCTGACTGA	TCTTCAAGAA	GGAGCAAGCC	ATGTGCGCTG	AGAAGATCCA	360
GAGGGCCCAT	GAGCTGATGG	GCTTCAATGA	TTCTCTTCCA	GGCTGTCTCT	GGATGTGGGA	420
CAACATCACG	TGTTGGAAGC	CGGCCCATGT	GGGTGAGATG	GTCTTGGTGA	GCNGCCCTGA	480
GCTCTTCCGA	ATCTTCAACC	CAGACCAAGT	CTGGGAGACC	GAAACCAATG	GAGAGTCTGA	540
TTTTGTGAC	AGTAATCTCT	TAGATCTCTC	AGACATGGGA	GTGGTGAGCC	GGAATGTCAC	600
GGAGGATGGC	TGGTCGGAAC	CCITCCCTCA	TTACTTTGAT	GCTGTGGGTT	TTGATGAATA	660
TGAATCTGAG	ACTGGGGACC	AGGATTATTA	CTACCTGTCA	GTGAGGCGCC	TCTACACGGT	720
TGGCTACAGC	ACTACCTCTG	TCACCCCTAC	CAGTGCATG	GTATCTCTTT	GTGCTTCCG	780
GAAGCTGCAC	TGCACACGCA	ACTTCATCCA	CATGAACCTG	TTTGTGTCTT	TGATGCTGAG	840
GGCGATCTCC	GTCTTCATCA	AAGACTGGAT	TCTGTATGCG	GAGCAGGACA	GCAACCACTG	900
CTTCATCTCC	ACTGTGGAAT	GTAAGGCCCT	CATGGTTTTC	TTCCACTACT	GTGTTGTGTC	960
CACTACTTTC	TGGCTGTCTA	TGGAGGGCCT	GTACCTCTTC	ACTCTGCTGG	TGGAGACCTT	1020
CTTCCCTGAA	AGGAGATACT	TCTACTGGTA	CACCATCATT	GGCTGGGGGA	CCCAACTGTT	1080
GTGTGTGACA	GTGTGGGCTA	CGCTGAGACT	CTACTTTGAT	GACACAGGCT	GCTGGGATAT	1140
GAATGACAGC	ACAGCTCTGT	GSTGGGTGAT	CAGAGGCCCT	GTGGTTGGCT	CTATCATGGT	1200
TAACCTTGTG	CTTTTATTG	GCAITATCST	CATCCTTGTG	CAGAACTTTC	AGTCTCCAGA	1260
CATGGGAGGC	AATGAGTCCA	GCAITACTTT	GCGACTGGCC	CGGTCCACCC	TGCTGCTCAT	1320
CCACTATTTC	GGATATCCCT	ACACAGTATT	TGCCTTCTCC	CCAGAGAATG	TACGCAAAAG	1380
GGAAAGACTC	GTGTTTGAGC	TGGGGCTGGG	CTCCTTCCAG	GGCTTTGTGG	TGGCTGTGCT	1440
CTACTGTGTT	CTGAATGGTG	AGGTACAAGC	GGAGATCAAG	CGAAATGTCG	GAGCTGGGAA	1500
GGTGAACCGT	TACTTGGCTG	TGGACTTCAA	GCACCGACAC	CGTCTCTGCG	CCAGCATGGT	1560
GSTGAATGGG	GGCACCAGC	TCTCCATCCT	GAGCAAGAGC	AGCTCCCAAA	TCCGATGCTC	1620
TGGCTCCCT	GCTGACAATC	TGGCCACCTG	AGCCATGCTC	CCCT		

Seq ID NO: 276 Protein sequence
Protein Accession #: NP_001109.1

1	11	21	31	41	51	
MAGVVHVSLA	AMCGACFWGR	GRLRRGRAAC	KSAQRHIGA	DLPLLSVGGQ	NCWPRSVMAG	60

5 VVHVSLAALL LLEPMAPAMHS DCIFPKKEQAM CLEKIQRANE LMGFNDSSPG CPGMMDNITC 120
 WKPAHVGENV LVSCPELEFRI FNPQVWRE TIGESDFGDS NSLDLSDMGV VSRNCTEDGW 180
 SEPFPHYFDA CGFDEYSEET GDQDYVYLSV KALYTVGYST SLVTLTAMV ILCRPRKLHC 240
 TRNFHMLNLF VSPMLRALSV FIKDWILYAE QDSNHCFIST VECKAVMVFF HCVVSNYFW 300
 LFIEGLYLFET LLVETFFPER RYFYWYTIIG WGTPTVCVTV WATRLYFDD TGCWDMMDST 360
 ALWWVKGFPV VGSIMVNFVL FIGIIVILVQ KLQSPDMGSG ESSYLRRLAR STLLILPLFG 420
 IHYTVFAFSP ENVSKRERLV FELGLGSEFG FVVAVLYCFL NGEVQAEIKR KWRSEKVNRY 480
 PAVDFKRRHP SLASSGVNGG TQLSILSKSS SQIRMSGPLA DNLAT

10 Seq ID NO: 277 DNA sequence
 Nucleic Acid Accession #: NM_004000.1
 Coding sequence: 36..1193

15 1 11 21 31 41 51
 AGAAGAAGCT GGCCAAGGAT ATGGGAGCAA CCACCATGGA CCAGAAGTCT CTCGGGCAG 60
 GTGTAGTGT CTGTCTGCTT CTCCAGGGAG GATCTGCCTA CAAACTGGTT TGCTACTTTA 120
 CCAACTGGTC CCAGGACCGG CAGGAACCGA GAAATTCAC CCCTGAGAAT ATTGACCCCT 180
 TCCTATGCTC TCATCTCATC TATTCATTGG CCAGCATCGA AAACAACAAG GTTATCATCA 240
 20 AGGACAAGAG TGAAGTGATG CTCTACCAGA CCATCAACAG TCTCAAAACC AAGAATCCCA 300
 AACTGAAATC TCTCTTGTCC ATTGGAGGGT ACCTGTTTGG TTCCAAAGGG TTCCACCTTA 360
 TGGTGGATTC TCTACATCA CGCTTGAAT TCATTAACCT CATAATCCTG TTTCTGAGGA 420
 ACCATAACTT TGATGGACTG GATGTAGCT GGATCTACCC AGATCAGAAA GAAACACTC 480
 25 ATTTCACTGT GCTGATTCT GAGTTAGCAG AAGCCCTTCA GAAGGACTTC ACAAAAATCA 540
 CCAAGGAAAG GCTTCTCTTG ACTGCGGGGG TATCTGCAGG GAGGCAATG ATTGATAACA 600
 GCTATCAAGT TGAGAAACTG GCAAAAGATC TGATTTTCAT CAACCTCTCTG TCCTTTGACT 660
 TCCATGGGTC TTGGGAAAGG CCCCTTATCA CTGGCCACAA CAGCCCTCTG AGCAAGGGGT 720
 GGCAGGACAG AGGGCCACAG TCCTACTACA ATGTGGATA TGCTGTGGGG TACTGGATAC 780
 30 ATAAGGGAAT GCCATCAGAG AAGGTGGTCA TGGGCATCCC CACATATGGG CACTCCCTTC 840
 CACTGCGCTC TGCAGAAACC ACCGTGGGGG CCCCCTGCTC TGGCCCTGGA GCTGCTGGAC 900
 CCATCAGAGA GTCTTCAGCG TTCTTGGCCT ATTATGAGAT CTGCCAGTTC CTGAAGGAG 960
 CCAAGATCAC GCGCTCCAG GATCAGCAGG TTCCCTACGC AGTCAAGGGG AACCAAGTGG 1020
 TGGGCTATGA TGATGTGAAG AGTATGGAGA CCAAGGTTCA GTTCTTAAAG AATTAAACC 1080
 35 TGGGAGGAGC CATGATCTGG TCTATTGACA TGGATGACTT CACTGGCAAA TCTTGCAACC 1140
 AGGGCCCTTA CCCTCTTGTG CAAGCAGTCA AGAGAAGCCT TGGCTCCTTG TGAAGGATTA 1200
 ACTTACAGAG AAGCGCCCAA GATGACCTTG CTGCTCTGGG CCGTCTCTCT CCCAGGATT 1260
 CTATGTTGGG ATTCCCTTTG CCAGGCTGGC CTTTGGATCT CTCTTCCAAG CCTTCTGTA 1320
 40 CTTCTCTCTA GATCATAGAT TGGACCTGGT TTGTGTTCC TGCAGCTGTT GACTTGTGTC 1380
 CCTGAAGTAC AATAAAAAAA ATTCATTTTG CTCCAGTA

Seq ID NO: 278 Protein sequence
 Protein Accession #: NP_003991.1

45 1 11 21 31 41 51
 MDQKSLWAGV VVLLLLQGGG AYKLVCTFTN NSQDRQEPGR FTRFNIDPFL CSHLIYSFAS 60
 IENKVLIKD KSEVMLYQTI NSLKTKNPKL KILLSIGYL FGSKGFBPMV DSSTSRLEFI 120
 NSIILFLRNH NFDGLDVSWI YPDQKENTHF TVLIHELAEA FQKDFPKSTK ERLLLTAGVS 180
 50 AGRQMDINSY QVEKLAKDLD FINLLSPDFH GSWEKPLITG HNSPLSKGNQ DRGPSSYINV 240
 EYAVGVWIHK QMPSSEKVVMG IPTYGHSTFL ASABTTVGAH ASGPBAAAGPI TESSGFLAYY 300
 ETCQFLKAGK IYRLQDQVVF YAVKGNQWVG YDDVKSMETK VQFLKNLNLG GAMNWSIDMD 360
 DFTGKSCNQG FYELVQAVKR ELGSL

55 Seq ID NO: 279 DNA sequence
 Nucleic Acid Accession #: NM_015166.1
 Coding sequence: 116..1249

60 1 11 21 31 41 51
 TGCCTGGAAGT CCTCTACCCA GAGACCACTG CTCCCAACGG CAGAGCAGOG GGGGAGATAA 60
 AGAAGTGGTG ACACGTGGCT GTACATTTCAG CACAGCTGTG GTGTCCCAA GTGCCATGAC 120
 CCAGGAGCCA TTCAAGAGAG AGCTGGCCTA TGACCGGATG CCAACGCTGG AGCGGGGCGG 180
 GCAAGACCCC GCCAGCTATG CCCCAGACGC GAAGCCGAGC GACCTGCAGC TGTCGAAGAG 240
 65 ACTGCCCCCG TGCTTCAGCC ACAAGACGTG GGTCTTCTCT GTGCTGATGG GGAGCTGCCT 300
 CCTGTGTACC TCGGGGTTTT CGCTGTACCT GGGGAACGTG TTCCCGGCTG AGATGGATTA 360
 CTTGCGCTGT GCTGCAGGCT CTGTGATCCC CTCCGCAATT GTGAGCTTCA CGTCTCCAG 420
 GAGGAACGCC AATGTGATTC CCAACTTTCA GATATTGTTT GTTTCACAGT TTGCTGTGAC 480
 CACTACGTGT TTAATTGGT TTGGATGCAA ACTAGTCTCT AADCCATCAG CAATAACAT 540
 70 CAACTTCAAC CTGATCTGCT TGCTCTGCTG GGAGCTGCTC ATGGCGGCA CGGTGATCAT 600
 CGCTGCACCG TCCAGCGAGG AGGACTGCAA GAAAAAGAAG GGCTCCATGT CTGACAGCGC 660
 CAACATTCTG GACGAAGTGC CATTTCTCTG TCGGTCCTG AATCTTACT CAGTCGTGCA 720
 GGTAAATCCA GGCATCTCTG CGGTCTCTGG GGGGATCATT GCGCTGAAGG TGGATGACTC 780
 AGTTCAGGCG CCACACCTCT CAGTGACGTT CTTTGGATC CTAGTGGCCT GCTTTCAGAG 840
 75 TGCCATTGCC AGTCATGTGG CAGCAGATG TCCAGCAAG TGCTGTGGTG AGGTCTGAT 900
 TGCCATAAGC AGCCTCACGT CTCGGCTGCT GTTCACAGCC TCTGGATATC TGTCTTTCAG 960
 CATCATGAGA ATGCTGAGGA TGTTTAAGGA TTACCGCCA GCCATAAAC CATCTACGA 1020
 TGTGCTGCTG CTGCTGCTGC TGCTAGTGCT CCTGCTGAG GCGCGCTTCA ACACGGGAC 1080
 CGCCATCCAG TGGTGTGGCT TCAAGGTCAG TGCAAGGCTG CAGGGTGCACT CCGGGACAC 1140
 80 CCAGAACGGC CCACAGAGC GCCTGGCTGG GAGGTGGGCC AAGAGCCCC TGAAGGAGT 1200
 GGACAGGAG AAAGCCTGGA GAGCCGTCGT GGTGCAATG GCCCAGTGAC CCCAGACGC 1260
 GGAAACCGGG CTGACAGGCG CAGCCTGGCC CCAAGCATGG AAAGGCACAA CCCCTAATGG 1320
 CCCTGAGCTA CTGCTCTTAA CACCTCTTTT CCGTGTGTG AGGGCAAAAC AGGCTGAGG 1380
 TGGGTTTTTC ACTTCCTAGG GTAGTTTAAT TTTAAATAG GCGAATGTTG GCTAGTCTGT 1440
 GCCTCAGTGA GATCAGTCAG CTCCGAGTGG CTCCCGTCTC GTAACAGCAG GAGCATGGCC 1500

5 GCAACTTCCC AGGCCGAGGA AGGGCCCCCG GCTCGGCCTC TTGAGAGCCC CACCCCTGAA 1560
 CTGGCCCCAG CTCCTCTTCC TGCCTCTCTC ATGGCTTGGG CTGGAGTGGG CTCTCTGGAC 1620
 CTGACCAGAC TGTGGGTCCC TGCCTCTCTC GCCACTCTG ACCGGGCTTC CTCCCTCCAC 1680
 GCTTAGGGTC TGTCCCGGGT ACTCAGTCAG CCCAGTGGGA TCTTACCAC TTCCCTGCAA 1740
 10 GGTSCACCTG CCCCAGGCTC AGGCTGCCCA GCGGCTCTTC CTGGACAGTG AGAGCAGGGC 1800
 TTGGCGCTTC TGTCTTGGCC GGGGAGCCGC AGGGGCCCTT CCTCCAGAGC CTGGCGCAA 1860
 GCACACAGG CTGCCGCTGC TCTCCAGGT GAAATCCACA CCAGTCCAGC CCGGTCGCC 1920
 TGCCCTGTCT CCTACTTAG ACCCAGTCAT TCTAGAGGA TCCACCGCCA CACTGGCCGG 1980
 CCGAAGTCTT GGTGTCTGTC ATGCCAGCT TGGAGTGGCA CGTGGCCGCT GCCACGTC 2040
 CCGGCACTGT CATGCCAGC TTGGAGTGGC ACATGGCCGC TGGCCACGTC CCGGGCACTG 2100
 TCATGCCAGC CTGGAGTGC CAGCTGGCCG CTGCTGTGAC AGGCAGTGT CTGGGGGTG 2160
 GGGCTGCATC CAGGCTTTG TAAACCGCT GGAACAGCTC TCCCTGGCCC CAGTGACCGG 2220
 GGGAACTGTA GCCCTCCCT CTCTGTGTG CTCCATTAC TCAAAATGCA GACAGATCA 2280
 15 GGTCAAGACC CAGGAATCT CACAGGTCA CCCAGCGCCC TCTACCTCT AGCAAGTACT 2340
 TTGTCTTGT CATGCCAGC AAGGCCCCAG GGCAGCGCTC TTCTCCATC CCGCTGTCTT 2400
 GGGGTCTTAG GTTACAGCCC AGGCGGTGAC TGCCACCTG CCAGGCTGCA GGGACAGTTG 2460
 GGTGTGAGAA TAACACTGCG TTTGGGTAGT GGCATGGCCA GAGTGGGT TCCCTGCGTC 2520
 TCTCTGTCCC GAGGGCGCTT GGGTCTCTCC AGCTGACGGC AGTAAATCCA CAGTGAGTTG 2580
 GGGCGACTGT GAAACTGGAA TGCTGTACT TTGATAATTA CTTCACGCA GGTGTCTTCC 2640
 20 TTCAAAATG TTTGTCTCT TCTCTTCTGA TCTGAGAAGA CATGAACGTT TTCTTTCAC 2700
 GCGCGTGGG TGTATTGACT GGTCCCCAT GGGCTGTGCG AAGGCCCGG AGATGCATCT 2760
 GTGGCTTGG GCCATCAAGA TCAAGAACC AGGAGGCTG GAGATGCGC CTGGATGGGG 2820
 CCGCTCTGAG ACCCTGCCG GGGGTGTGAG GACCTCTCCA GGTTCGCCAC TGGGAACAG 2880
 GAGTGACTCT GCTCTGCAAG ATACCTTCTAT GGTGTTCATG ACAAGTGGAA TCATTATTTT 2940
 25 CAACCAATGA AGGGGATGC AGGCAAGACA CCTTCCGAGC TGCTCCTAGA GGGGACAGC 3000
 CAGGCCCTCT CTGCACTCTT CCGCAGCTCC GGAAGGACAC AGTCAGGGGC CCGGCAACA 3060
 CTTTGGCCAC ACCCCCAAC AAGCGCCACC GTGGGAGAGG AGAGGCTGCT GTCACTGCTA 3120
 CCGGATGACG ACCCCACCT GTCTGCAAGC CACCTCCACC TCCCTGCGAG TTTGAGGCTG 3180
 CCGGGGTCTG CTCTCTGGGA TGGGGTGGGA GCCACAGGGA CGACCCGGGG CCGGCTGATG 3240
 30 TCTTCTTGG GTCAGACCTG AGAGCTCAAG TTTCAAGTCA AGAATTAGGC ACTTGAACG 3300
 TTTTGTCTG CTGCACTCTT CTATTCTCT TATTTTAGAG CGCTTAAAA ATCCGGAAAA 3360
 ATGGGGTTA AAGAAGCTGT CTCTTTCAGT CTACATTTTT GTTAAATACG CTGAGCAAT 3420
 AAGGCTGAC TTGCAGAGT G

Seq ID NO: 280 Protein sequence
 Protein Accession #: NP_055981.1

1 11 21 31 41 51
 40 MTQEPFREEL AYDRMPTLER GRQDPASYAP DAKPSDLQLS KRLPFCFSHK TWVFSVLMGS 60
 CLLVTSQFSL YLGNVFAEM DYLRCAAGSC IPSAIVSETV SRRNANVIPN FQILFVSTFA 120
 VTTTCLINWF CKLVNBSAI NINFLILL LLELMBATV IIAARSSEED CKKKRGSMSD 180
 SANTLDEVPF PARVLKSYSV VSVIAGISAV LGGLIALNVD DSVSGPHLSV TFFWILVACF 240
 45 PSAIAHVAA BCPSKCLIVEV LIAISLSTSP LIPTASGYLS FSIIMRIVEMF KDYPFAIKPS 300
 YDVLILLILL LLLVLLQGLNT GTAIQCVRFK VSARLQGASH DTQNGPQERI AGEVARSPLK 360
 EFDKKEAWRA VVVQMAQ

Seq ID NO: 281 DNA sequence
 Nucleic Acid Accession #: NM_004518.1
 Coding sequence: 43..2577

1 11 21 31 41 51
 55 GCTGAGCCTG AGCCCGACCC GGGGCGCTTC CGCCAGGCA CCATGGTGC AAGTCCGCG 60
 AACGCGCGCG TATACCCCTG CCGGAGCGGG GAGAAGAAGC TGAAGTGGG CTTCGTGGGG 120
 CTGACCCCGG GCGCGCCCGA CTCACCCCGG GACCGGGCGC TGCTGATCGC CCGCTCCGAG 180
 GCCCCCAAGC GCGGACGAC CCTCAGCAAA CCTCGCGCGG GCGCGCGCGG CCGCGGGAAG 240
 CCCCCAAGC GCAACGCTT CTACCGCAAG CTGCAGAAAT TCCCTACAA CGTGCTGGAG 300
 CCGCGCGCGG GCTGGGCTT CATCTACCAC GCTACGCTGT TCCTCTCTGT TTTCTCTG 360
 60 CTCTGCTGT CTGTGTTTC CACCATCAAG GASTATGAGA AGAGCTCGGA GGGGGCCCTC 420
 TACATCTGAG AATCGTGAC TATCGTGGT TTTGGCGTGG AGTACTCGT GCGGATCTGG 480
 GCGGCAAGCT GCTGCTGCC GTACCGTGGC TGGAGGGGGC GGTCAAGTT TGCCCGGAAA 540
 CGGTTCTGTG TGATTGACAT CATGGTCTC ATCGCTCCA TTGCGGTGCT GGCAGCCGCG 600
 TCCAGGGCA ACCTCTTTGC CATATCTGCG CTCGAGAGCC TGCGCTTCT GCAGATTCTG 660
 65 CGGATGATCC GCATGGAAG GCGGGGAGGC ACCTGAAGC TGCTGGGCTC TGTGCTCTAT 720
 GCCACAGCA AGGAGCTGCT CACTGCTGAG TACATCGGCT TCCTTTGTCT CATCTCGGCC 780
 TCGTCTCTGG TGTACTTGGC AGAGAAGGGG GAGAAGGACC ACTTTGACAC CTACGCGGAT 840
 GCACTCTGCT GGGGCTGAT CACGCTGACC ACCATTGGCT ACGGGACAA GTACCCCCAG 900
 ACTTGAAGC GCAGGCTCTT TGGGGCAACC TTCACCTCA TCGGTGTCT CTCTCTCGCG 960
 70 CTGCTCTGAG GCATCTTGG GTCTGGGTTT GCGCTGAAGG TTCAGGAGCA GCACAGGCAG 1020
 AAGCACTTTG AGAAGAGGCG GAACCCGCGA CGAGGCTGTA TCCAGTGGC CTGGAGATTG 1080
 TACGCCACCA ACCTCTCGCG CACAGACTG CACTCCAGT GGCAGTACT CAGCGGAAG 1140
 GTCACCTGCG CCATGTACAG ACTTATCCCC CCGCTGAACC AGCTGGAGCT GCTGAGGAAC 1200
 CTCAGAGATA AATCTGGACT CGCTTTGAGG AAGGACCCCG CCGCGAGGCC GTCTCCAAGC 1260
 75 CAGAAAGTGA GTTTGAAGA TCGTGTCTTC TCCAGCCCGG GAGGCGTGGG TGCCAAAGGGG 1320
 AAGGGTCTCC CCGAGGCCCA GACTGTGAGG CGGTACCCA GCGCGGACCA GAGCCTCGAG 1380
 GACAGCTCCA GCAAGCTGCC CAGAGCTGAG AGCTTCGGGG ACCGCGAGCG GGCAGCGGAG 1440
 GCTTTCCGCA TCAAGGTGCG CGCTCACCG CAGAACTCAG AAGAAGCAAG CCTCCCCGGA 1500
 GAGGACATTG TGGATGACAA GAGCTGCCCC TCGAGATTG TGACGAGGAG CCTGACCCCG 1560
 80 GGCCTCAAGT TCAGATCAG AGCCGTGTGT GTCTGCGGT TCCTGGTCTC CAAGCGGAAG 1620
 TTCAGGAGA GCTCTCGGCC CTAAGAGCTG ATGGAAGTCA TCGAGCAGTA CTCAGCCGCG 1680
 CACTTGGAGA TGCTGTCCCG AATTAGAGC CTGCAGTCCA GAGTGGACCA GATCGTGGGG 1740
 CCGGGCCCGG CGATCAGCGA CAGGAGCCCG ACCAAGGGCG CCGCGAGGCG GAGCTGCCC 1800
 GAGGACCCCA GCATGATGGG ACGGCTCGGG AAGGTGGAGA AGCAGGTCTT GTCCATGGAG 1860

1061

CGTGTGCCCG CCATGCCCGC GCCCAAGCC ATTGCAGTCT TCCATCCTCT GGCCTGACG 6960
 GTGGCTGCAG CTTCCCAATT TGCGCGGTG CCTCTGGCTG TCTGCACCTT TGTTCATGCT 7020
 CCAAGAGACA TTTCATAATG CTTTCAGTAC CAGCGTACAC TCTGACCAT TTTGTATGTG 7080
 TCCTTGTGCC GTAGTGACCA GGCCTTTTT TGGTGGATGT GTTACCCCGC ACACCTCAAT 7140
 CTCACCTTTC TGCACGTCOC ATTTCTTAGG GATAGACGCC CAGGGAATGA ACTCTAGTTC 7200
 TCTAACAGAT TAGCTGAGAT ATTAACCTAC TCACACGGAC AGGTTGATGC CAGAGCCGTA 7260
 AGAATGCGCC AGTGGCGGTT TGCGGGGGAC TTCGGGTGTG GGGTCTCTGC GCCCGGATGG 7320
 CCGTGGAGAG TCTGGGGAT CCTGCTGCC ACGGGGACGA GTTCGGACGC CAGGTGGACC 7380
 TGTGCACTCA GTAAACGCA GTGATTCAAA AAAAAAAAAA

'Seq ID NO: 282 Protein sequence
 Protein Accession #: NP_004509.1

1 11 21 31 41 51
 MVQKSRNGGV YPGPSGEKKL KVGPFVGLDPG APDSTRDGL LIAGSEAPKR GSILSKFRAG 60
 GAGAGKPPKR NAFYRKLQNF LYNVLERPRG WAFYHAYVF LLVPSCLVLIS VPSTIKEYEK 120
 SSBGALYILE IVTVVVFVE YFVRINAAGC CCRYRGMWRGR LKFAKPFPCV IDIMVLIASI 180
 AVLAAGSQGN VFATBALRSL RFLQILRMIR MDRRGSTWKL LGSVVYABSK ELVTAWYIGF 240
 LCLILASFVN YLAAGENDH FDTYADALWW GLITLTITIGY GDKYPQTNWG RLLAATFTLI 300
 GVSFFALPAG ILGSGFALKV QEQHRQKHE KRNNEAAGLI QSAWRPYATN LSRDTLHSTW 360
 QYYRTVTVP MYRLIPPLNQ LELLRLNLSK SGLAFRKDPF PEPSPSQKVS LKDEVFSSPR 420
 GVAAGKGSFP DAKTAGCFAA CDTAGCAAAG GATCTGGGAC TAAGGGTAGA GGAACCTGCC 480
 BASLPGEDIV DDKSCPCFEV TEDLTPLKLV SIRAVCYMRF LVSKRKPKES LRPFYDMVDVI 540
 EQYSAGHLDL LSRIKSLQSR VDLVGRGPA ITDKORTKGP ASAELEFEDPS MMGRLEGVEK 600
 QVLSMEKKLD FLVNIYMORM GIPPTETZAY FGAKPEPEAP PYHSPEDSRE HVDREGCIVK 660
 IVERSSTGTG KNFSAPPAAP PVQCPPTSM QEQSHPRQGH GTSFVCDGEGS LVRIPPPPAH 720
 ERSLAYGGG NRASMEFLRQ EDTPGCRPPE GTLRDSDTSI SIPSVIHEDL ERSFSGFSIS 780
 QSKENLDALN SCYAAVAPCA KVRPYIABGE SDTDSDLCTP CGPPPSRATG EGPFQDVGWA 840
 GPRK

Seq ID NO: 2838 DNA sequence
 Nucleic Acid Accession #: AF152496.1
 Coding sequence: 1..2391

1 11 21 31 41 51
 ATGGAGGCGG GAGGAGAGCG ATTTCTTAGA CAAAGGCAAG TCTTGCTCTC CTTTGTITTT 60
 CTGGGAGGGT CTCCTGGCTGG GTCCGAGTCA AGACGCTATT CTCTGGCTGA GGAAAAAGAG 120
 AAGGCTTTT TAATAGCAA CDTAGCAAAG GATCTGGGAC TAAGGGTAGA GGAACCTGCC 180
 GCGAGGGGGG CCCAAGTTGT GTCCAAAGGG AACAAACAGC ATTTCTAGCT CAGTCATCAG 240
 ACAGGTGATT TGCTCTCTAA TGAGAAATTG GACCGGGAGG AGCTATGCGG CCCACAGAA 300
 CCATGCATAC TACATTTTCA GATATTACTG CAAAACCTTT TGCAATTCGT TACAAACGAG 360
 CTCCTGATCA TAGATTAATA TGACCATCTC CCGGTATTCT TTGAAATGA AATGCATCTG 420
 AANAATCCTAG AAGGACTCTC GCAGGAACA GTAATTCCTT TGGGAATATG TGAGGACTTG 480
 GATGTGGGAA GAACAGCCTC CAAAACCTAC ACTATCACTC CGAATTCCTA CTTCCACGTA 540
 CTCACCTGCA GTGCTAGGGA CGGAAGGAAG TACCGGAAAC TAGTACTGGA TAAGGCTCTC 600
 GATCGGGAGG AGCAGCGCGA ACTCAGCTTA ACCTCTACCG CGCTGGACGG CGGCTCTCCC 660
 CCTCGGTCTG GGCAGGCCCA GATAAACATC CAGGTCTTAG ATATAAACA CAATGACCA 720
 GAATTGTGAC AGCGCTCTA TGAGGTGCA GTTCTAGAGA ATACCCCGGT TAACTCTGTC 780
 ATTGTCACTG TCTCGCTCTC TGACTTAGAT ACAGGAAGTT TTGGGACAA ATCATATGCA 840
 TTTTTCATG CTTCTGAAGA AATTGCGAAA ACTTTTCAGC TAAATCCAAT TACTGGTGAT 900
 ATGCAACTGG TCAATATTTT GAATTTTGA GOGATTAAAT GTTATGAAGT CGACATCGAG 960
 GCGAGGATG CCGAGGGCCT ATCCGGAAG TCTACAGTCA TAGTCCAGGT GGTGTATGTC 1020
 AAGGACAAAC CACGGAACCT GACCTTGACT TCAGTAAACA GCOCTATTC TGAAGAACCTG 1080
 GGAGAGACTG TACTGCTGT TTTCACTGTT TCTGATCTAG ACTCTGGAGA CAACGGAAGA 1140
 GTGATGTGTT CCATTGAGAA CAATCTCCCC TTCTTCTCTA AACCATCTGT AGAGAAATTT 1200
 TACACCTAG TGTACAGAGG CGGCTGGAC ABAGAGACCA GATCCGAGTA CAACATTACC 1260
 ATCACTATCA CTGACCTGGG GACACCCAGG CTGAAACCA AGTACAACAT AACCGTGCTG 1320
 GTCTCCGAG TCAATGACAA CGCCCCCACC TTCAACCAAA TCTCTACAC CCTGTGCTC 1380
 CGCGAGACA ACAGCCCCGC CCGCACATC GGCAGTGTCA GCGCCACAGA CAGAGACTCA 1440
 GGCACCAACG CCGAGGTAA CTACTGCTG CTGCGCCGCC AGGACCCGCA CCGCCCTC 1500
 TCTTCCCTGG TCTCATCAA CGCGGACAA GGCACCTGT TTGCCCCAG GTCCTGAC 1560
 TAGAGGCCCT TGACGCGTT CAGGTCCCG GTGGGCGCA CAGACCTGG CTCCCGGCT 1620
 TTGAGCAGG AGCGCTGGT GCGCTGCTG GTCCTGGAC CCAACGACAA CTCGCCCTC 1680
 GTCTGTATCC CGCTGCAGAA CGGCTCCCG CCTGTACCG AGCTGGTCC CCGGGCGGCT 1740
 GAGCGGGGCT ACCTGGTGAC CAAGGTGGT GCGGTGGAC GCGACTCGGG CCGAACGCC 1800
 TGGCTGTGTT ACCAGCTGCT CAAGGCCAG GAGCCCGGG TGTTCGGCT GTGGGCGCAC 1860
 AATGGCGAAG TGCGCACCGC CAGCTGCTG AGCGAGCGG ACBCGCGCA GCACAGGCTG 1920
 GTGGTGTGTC TCAAGGACAA TGGCGAGCT CCGCGCTCG CCAACGCGAC GCTGCATGTG 1980
 CTCCTGTGTC ACGGCTTCTC CCGGCCCTAC CTGCTCTCC CGGAGGCGGC ACGGCGCCAG 2040
 GCCCAGGCCG ACTTGTCTAC GGTCTACTTG GTGGTGGCA TGGCTCGGT GTCTTCGCTC 2100
 TTCTCTTTT CGGTGCTCTT GTTCTGTCG GTGCGGCTGT GCAGAGGAG CAGGGCGGCC 2160
 TGGTGGGTC GCTGCTCGGT GCGCGAGGCG CCTTTCCAG GGCAGATGTT GGAATGAGC 2220
 GGCACCGGA CCGCTGCCCA GAGCTACAG TACGAGGTGT GTCTGACTGG AGGCTCCGG 2280
 ACAATAGAT TCAAGTTCTT GAAGCCAAT ATCCCAACT TCGTTGCTCA GGTGCGAGG 2340
 AGGGTTAGCG AGGCAATCC CAGTTTCAGG AAGAGCTTTG AATTCACTTA A

Seq ID NO: 284 Protein sequence
 Protein Accession #: AAD43757.1

1 11 21 31 41 51
 MEAGGERFLR QRQLLLFVF LQSLAGSES RRYSVABERE KGFLIANLAK DLGLRVEELA 60

5 ARGAAQVVSKEG NKQHFQLSHQ TGDLLNEKL DREELCGPIE PCILHFQILL QNPLQFVTNE 120
 LRLIDVNDHS PVFFENEMHL KILESTLPOT VIPLGNAEDL DVGNSLQNY TITFNSHPHV 180
 LTRRRDGRK YPELVLDKAL DREEQPELSL TLTAIDGGSP PRSGTAQINI QVLIDINDAP 240
 EFAQPLYSVA VLENTFVNSV IVTVSASLSD TGSFGTISYA FFHASEEIRK TFQLNIPITGD 300
 MQLVKYLNF EAINSYEVIE AKDGGGLSGK STVIVQVVDV NDNPPETLS SVNSPIFENS 360
 GETVLAVFSV SDLSGSDNGR VMCSIENLPL PFLKPSVENF YTLVSEGLD RETRSEYMIT 420
 ITITDLGTER LKTKYNITVL VSDVNDNAPA FTQISYTLFV RENNSPALHI GSVSATDRDS 480
 GTNAQVTVSL LPPQDPHPL SSVLSINADN GHLFALRSLD YRALQAFEFV VGATDRGSPA 540
 LSSEALVRVL VLDANDNSPF VLYPLQNGSA PCTELVPRAA EPGYLVTKV AVDGDGSGNA 600
 WLSYQLKAT EFGLFVVAH NGEVTRARLL SERDAKHRL VVLVDNGEP PRSATATLHV 660
 LLVDGFSQPY LPLPEAAPAQ AQADLLTVYL VVALASVSSL FLFSVLLFVA VRLCRRSRAA 720
 SVGRCSVPB3 PFPQGMVDVS GTGTLQSQSY YEVCLTGGSG TNEFKFLKPI IPNFVAQGA 780
 RVSEANPSFR KSFEFS

15 Seq ID NO: 285 DNA sequence
 Nucleic Acid Accession #: NM_001794.2
 Coding sequence: 15..2765

20 1 11 21 31 41 51
 CGCCGCGCGG GAAGATGACC GCGGCGCGCG GCGTGTCTCT TCTGTCTCTC TCGTCTCTCG 60
 GCGCCTCTCG GGCCTCAAT GAGGATCTTA CAAGTAGAGA GACCTGCAAG CCGTGGTCTT 120
 CTGAAGATGA TTACACGGCA TTAATCTCCC AAAATATTCT AGAAGGGGAA AAGCTACTTC 180
 AAGTCAAGTT CAGCAGCTGT GTGGGGACCA AGGGGACACA ATATGAGACC AACAGCATGG 240
 25 ACTTCAAGT TGGGGCAGAT GGGACAGTCT TCGCCACCGG GAGAGCTGAG GTCCCTCTCG 300
 AGCAGGTGGC GTTCACTGTG ACTGCATGG ACAGCCAGAC AGCAGAGAAA TGGGACGCGG 360
 TGGTGGCGTT GCTGTGTGGC CAGACCTCTT CCGCGCACTC TGGACACAAG CCGCAGAAAG 420
 GAAAGAAGGT CGTGTCTCTG GACCTCTCTG GCGCTCGGAA GGCACCTCTG CTGCGGTGGC 480
 30 CCGACGACCA GACAGCCGAC GGGCTGAGGC GCGCGAAACG GCACTGGGTC ATCCCGCCCA 540
 TCAACGTGCC CGAGACTCG CCGCGGCGCT TCCCGCAGCA GCTCGTGGG ATCCGCTCGG 600
 ACAAGACAA TGACATCCCC ATCCGCTACA GCATCAAGGG AGTGGCGCGC GACCAAGCCC 660
 CCAATGAGGT CTTAGCAATT GACTCCATGT CCGGCGGAT GTACGTACA AGSCCAATGG 720
 ACCGGGAGGA GCACGCTCTT TACCACTCTC GAGCCACGCG TGTGGACATG AATGGCAACA 780
 35 AGGTGAGGAA CCGCATGAGC CTGTACATCT ACCTCATCGA CATGATGAC AACCGCCCTG 840
 AGTTCATCAA CAGGTCTTAC AACGCTCCG TGGACGAGGG CTCCAAGCCA GGCACCTTAC 900
 TGATGACCGT CACGCGCAAC GATGCTGACG ACAGCACCAC GGCACACGGG ATGGTGGGGT 960
 ACCGATGCTT GACCCGAGCC CCACAGAGCC CGTCCCAAG TATGTTTACC ATCAACAGCG 1020
 40 AGACTGGAGA TATGTCACA GTGGCGGCTG GCTTGGACCG AGAGAAAGTT CAGCAGTACA 1080
 CAGTCATCTT TCAGGCGACA GATATGGAAG GAAATCTCAA CTATGGCCTC TCAACACAG 1140
 CCACAGCATC CATCAAGGTG ACAGATGTGA ATGACAAACC GCCAGAAATT ACCCGCAGCA 1200
 CGTTCGAGG GAGAGTCCCC GAAACCCCG TGGAGACCGT GGTGCGAAAC CTCACGCTGA 1260
 TGGACCGAGA TCAGCCCCAC TCTCCAAACT GGAATGCCGT TTACCGCATC ATCAGTGGGG 1320
 45 ATCCATCCCG GCACCTCAGC GTCCGACAG ACCCGTAAC CAACGAGGCG ATGGTCAACG 1380
 TGGTGAAGGC AGTGGACTAC GAGCTCAACA GAGCTTTTAT GCTGACAGTG ATGGTGTCCA 1440
 ACCAGGCGCC CCGTGGCGAG GGAATCCAGA TGTCTCTCCA GTCCAGGCA GGGGTGACCA 1500
 TCTCCATCAT GGCATCAAC GAGGCTCCCT ACTTCCCTC AAACCAAG CTGATCCGCG 1560
 TGGAGGAGGG CGTCCCCCCC GGCACCGTTC TGACCACTTT TTCACTGTG GACCTGAGC 1620
 50 GGTTCATGCA GCAGGCTGTG AGATACTCAA AGCTGTGAGA CCCAGGAGC TGGCTGCACA 1680
 TCAATGCCAC CAACGCGCAG ATCACCACGG CCGCAGTGCT GGCAGGTGAG TCCCTCTACA 1740
 CCAGAAACAA CGTCTACGAG GCCACCTTCC TGGCAGCTGA CAATGGGATA CCCCAGGCA 1800
 GCGGCAACCG GACCTTCAG ATCTATCTCA TTGACATCAA CGACAAGGCC CTGAGCTGCG 1860
 TGCCCAAGGA GGCAGAGATC TGGCAGAGC CCAACCTGAA CGCATCAAC ATCAGGCGCG 1920
 CCGACGCTGA CGTGCACCCC AACATCGGCC CCTACGCTCT CGAGCTGCCC TTTGTCCCGG 1980
 55 CCGCGGTGCG GAGAGACTGG ACCATCACCC GCGTGAACGG TGACTATGCC CAAGTCACT 2040
 TGGCATCTCT GTACCTGGAG GCGGGGATGT ATGACGTCCC CATCATGCT ACAGACTCTG 2100
 GAAACCCCTC CCGTCCAAAC AGTCCATCA TCAAGTCAA GGTGTGCCCA TGTGATGACA 2160
 ACGGGGACTG CACTACCAT GGCAGCTGG CAGCGGCTGG TCTGGGACCG GGTGCATCG 2220
 TGGCCATCTT CATCTGCATC CTCATCTGCG TGACCATGGT CCGTCTGTTT GTCATGTGGA 2280
 60 TGAACCGCGG AGAGAGAGAG CGCCACAGGA AGCAGCTGCT CATGACCCC GAGGACGAGG 2340
 TCGCGACAAA CATCTCAAG TATGACGAGG AAGGCGGTGG CGAGGAGGAC CAGGACTACG 2400
 ACCTCAGCCA CGTGCAGAG CCGGAAGCCA TGGGGACAGT GCCAAGCAA GCCCTGCGCG 2460
 TGCTTCGCTT GATGAGAGCG CCGTGGGCG CTGAGCCCCA GTACCCGATC AGGCCATGG 2520
 TGGCGCAACC AGGCGACATC GGTGACTTCA TCAATGAGGG ACTCCGCGCT GCTGACAAAG 2580
 65 ACCCCACGCG ACCCCCCAT GACTCCCTGC TGGTCTTCCA CTACGAGGGG AGCGGCTCCA 2640
 CCGCAGGCTC CGTCACTCTC CTGAACTCAT CCACTTCCGG GGCACCAAGC TACGATTACC 2700
 TCAACGACTG GGGGCGCAGA TTCAAGAAGC TGGCGGACAT GTATGAGAGT GGTGAAGAGG 2760
 ATTGACTGAC CTCGACCTT CGGACTGAGG TGAGAGCGGT GCTCGGAGCG CGGAGGAGCA 2820
 GGAAGTGAAG GAGGCGGCGG GTCTTCCCGA CTCCCTGCGG CTGTGTCTTT AGTGTCTTTA 2880
 70 GAGAGCCCCC CAATCCCCAC GTTGACTGT CTAGCATGAG CACCCACCCG CACAGCGCCC 2940
 TGCACCGCGC CGCTGCCAG CACCGCGCTG CCTGGCACTG AAGGACAGCA AGAGGCACTC 3000
 TGTCTTCACT TGAATTTCTT AGAACAGAG CACTGTTTTT AAAAAAAAAA AAAAAAAAAA 3060
 AAG

75 Seq ID NO: 286 Protein sequence
 Protein Accession #: NP_001785.2

80 1 11 21 31 41 51
 MTAGAGVILL LLSLSGALRA HNEELTTRET CKAGFSEDDY TALISQNIIE GKLLQVKFS 60
 SCVGTETGTY STNSMDFKVG ADGTVFATRE LQVPSQVAF TTTAWDSQTA EKNDVAVELL 120
 VAQTSSPHSG HEPQKQKVV ALDPSPPPKD TLLPWPQHQN ANGLRRRKRK WVIPPINVE 180
 NSRGPPFQQL VRIKDKND IPRTYSITGV GADQPFMEVF SIDSMSGRMY VTRPMREEH 240
 ASYHLRAHAV DMNGNKVEMP IDLYIYVIDM NDMRPEFINQ VYNGSVDEGS KPGTYVMIVT 300
 ANDADDSTTA NGMVRYRIVT QTPQSPSQNM FTINSETGDI VTVAAGLDRE KVQYTVIVQ 360

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

```

ATDMEGNLNY GLENTATAII TVTDVNDNPF EFTASTFAGE VPENRVRTV ANLTVMDRDQ 420
PHSPNNAVY RIISGDPSGH FSVRTDFVTN EGMVTVVKAV DYELNRAFML TVNVSNGAPL 480
ASGIQMSFQS TAGVTIISIND INEAPYFFSN HKLIRLEEGV PPGTVLITFS AVDPDRFMQQ 540
AVRYSKLSDP ASNLHNATN GQITTAVALD RESLYTKNNV YEATFLAADM GIPPASGTGT 600
LQIYLIDIND NAEPLLEKEA QICEKPNLNA INITAADADV DPNIGPYVFE LPFVPAAVRK 660
NWTITRLNGD YAQLSLRILY LEAGMYDVEI IVTDSGNPPL SNTSIIKVKV CPCDDNGDCT 720
TIGAVAAAGL GTGAIIVAIL CILILLTMVL LFMVMMKRRE KSRHTKQLLI DPEDDVRDNI 780
LKYDEFGGGE EDQDYDLSQL QQPEAMGHVP SKAPGVRRVD ERPVGAEPQY FIRPMVPHPG 840
DIGDFINEGL RAADNDPTAP PYDSLLVFDY EGSGSTAGSV SSLNSSSSGD QDYDYINDWG 900
PRPKLADMY GGGEED
  
```

Seq ID NO: 287 DNA sequence
 Nucleic Acid Accession #: AF152495.1
 Coding sequence: 1..2397

15
 20
 25
 30
 35
 40
 45
 50
 55
 60

```

1 11 21 31 41 51
| | | | |
ATGGAGGCGG GAGAGGGGAA GGAGCGCGTT CGAAACAAA GGCAAGTCCT GATATTCCTT 60
GTTTGTGCTG GCATAGCTCA GGCTAGTTGC CAGCCTAGGC ACTATTCAGT GGCCGAGGAA 120
ACGGAGAGTG GCTCCTTTGT GGCCAATTTC TTAAGAGACC TGGGCTGGA GATAGGAGAA 180
CTTGCTGTGA GGGGGGCCAG GGTGCTTTCC AAAGGAAAAA AATGTCATT GCAGTTCGAT 240
AGGCAGACCG GGGATTTGTT GTTAAATGAG AATTGGACC GGGAGGAGCT GTCCGGCCCC 300
ACAGAGCCCT GTGTCCCTACC TTTCAGGTG TTAGTAGAAA ATCCCTTGCA GTTTTTCAG 360
CGGGAGCTAC GGATTCAGGA CGTAAATGAT CATTCCCTAG TTTCCCTAGA CAAAGAAATA 420
CTTTTGAAAA TTCAGAAAG TATCACTCCT GGAACCTACT TCTTAATAGA ACGTGCCTAG 480
GACTTGGATG TAGGAACCAA CAGTCTCCAA AATTACACAA TCAGTCCCAA TTTCCACTTT 540
CATCTTAATT TACAGACAG TCTCGATGGC ATAATATTAC CACAGCTGGT GCTGAACAGA 600
GCCCTGGATC GCGAGGAGCA GCCTGAGATC AGGTTAACCC TCACAGCGCT AGATGGCGGG 660
AGTCCACCCA GGTCCGGCAC GGCCTCGGTA CGGATGGAAG TTGTGGACAT CAATGACAC 720
GTCCAGAGAT TTGCAAGCT GCTCATGAG GTGCAGATCC CGGAGGACAG CCGCTTGA 780
TCCAGGTTG CCATCGTCTC TGCAGGGAT TTAGACATTG GAACATAAG AGAAATATCT 840
TATGCAATTT CCCAGCATC TGAAGACATT CGCAAAACGT TTCGATTAG TGCAAAATCG 900
GGAGAACTGC TTTAAGACA GAAACTGGAT TTCGAATCCA TCCAGACATA CACAGTAAAT 960
ATTCAGCGCA CAGATGGTGG GGCCTATCTT GGAACCTGTG TGTATTGTGT CCAAGTGATG 1020
GATTTGAATG ACAATCTCC GGAATAACT ATGTGACAC TTATCAATCA GATCCAGAA 1080
AACTTGCAGG ACAOCTCAT TGCATGATC AGCGTTTCAG ATCCTGACAT CGGAGACAC 1140
GGAAGGATGG TGTGCTCAT CCAAGATGAT CTTCCTTTT TCTTGAAACC TTCGTTGAG 1200
AACTTTTACA CTCTTTGAT AAGCACGGCC CTGGACCGGG AGACAGATC CAGATACAC 1260
ATCACCATCA CCGTCACGCA CTTCGGGACA CCCAGGCTGA AAACCGAGCA CAACATAACC 1320
GTGCTGGTCT TGTAGGTCAA TGACAACGCC CCGGCTTCA CCAAAACCTC CTACACCTCG 1380
TTGCTCGCGG AGAACAACAG CCGCGCCCTG CACATCGGCA GGTTCAGGCG CACAGACAGA 1440
GACTCGGGCA COAACGCCCA GGTCACTTAC TCGCTGCTGC CGCCCGAGGA CCGCACCTG 1500
CCCTCGCTCT CCCTGCTCTC CATCAACGCG GACAACGCC ACCTGTGCG TCTCCAGTGG 1560
CTGACTACCG AGGCGCTGCA GCGCTTCGAG TTCGCGTGG GCGCCGAGA CCGCGGCTCC 1620
CGGCGGTTGA CCGAGCGAGC GTGCTGCTGC TGGACGCCAA CGACAACCTG 1680
CCCTCGTGC TGTACCGCT CGACAACGCG TCGCGCGCT GCACCGAGCT GGTGCCCGG 1740
CGGCGCGAGC CGGCTACTCT GGTGACCAAG GTGGTGGCG TGGACGCCGA CTCGGGCCAG 1800
AAGCGCTGGC TGTCTGATCA GCTGCTCAAG GCGACGGAGC CGGGCTGTT CGGCTGTGG 1860
GCGCACATG GCGAGGTGCG CACCGCCAGG CTGCTGAGGG AGCGCGAGCG TGCCAGACAG 1920
AGGCTGGTGG TGTCTGTCAA GGACAATGCG GAGCCTCCGC GCTCGGCCAC CGCCACGCTG 1980
CACTGCTGCC TGTGAGACCG CTCTCCCGAG CCTACTGTC TGTCCCGGA GCGGCGACCG 2040
GCCGAGGCCC AGGCGGACTT GCTCACCTTC TACCTGCTGG TGGCGTTGGC CTGGGTGCT 2100
TGTCTCTTCC TCTTCTCGGT GCTCCTGTTT GTGGCGGTGC GGCTGTGCG GAGGAGCGGG 2160
GCGGCGTGG TGGGTGCTG CTCGCTGCC GAGGCGGCTT TCCAGGGCA GATGTTGAC 2220
GTAGAGCGCA CCGAGGACCT GTCCAGAGC TACCACTAG AGGTGTGCT GACTGGAGGC 2280
TCCGGGACAA ATGAGTTCAA GTTCTGAAG CCAATTATCC CCAACTTGT TGTGAGGGT 2340
GCAGAGAGGG TTAGCGAGGC AATCCAGT TTCAGGAGA GCTTTGAAT CACTTAA
  
```

60
 65
 70
 75
 80

```

Seq ID NO: 288 Protein sequence
Protein Accession #: AAD43756.1

1 11 21 31 41 51
| | | | |
MRAGEGKERV PKQRQVLIF VLLGIAQASC QPRHYSVAEE TEGGSFVANL LKDLGLEIGE 60
LAVRGARVVS KQKQMLQFD RQTGDLLINE KLDREELCGP TEPCLVLPQV LLENPLQFFQ 120
AELRIIRDND HSPVFLDKBI LKIPESITP GTTFLIERAQ DLDVGTNSLQ NYTISENPHF 180
HNLQDSLDG IILPQLVLNR ALDRHEQPEI RLTLTALDGG SPFRSGTALV RIEVDININ 240
VPEFANGLYE VQIPEDSPVG SQVAIVSARD LDIGTINGEIS YAFSQASEDI RXTFRLSAKS 300
GELLRLQKLD PESIQTYTVM IQATCGGLS GTCVVFVQVM DLNDMPBELT MSTLINQIPE 360
MLQDTLIAVF SVSDPDSGDN GRMVCISQDD LFFFLKPSVE NFFTLVISTA LDRETRSEYN 420
ITITVTDPGT FRLTEENIT VLVSDVNDNA PAFTQTSYTL FVRENNSPAL HIGSVSATDR 480
DSGTHAQVTY SLLPFPQDEHL PLASLVISINA DNGHLFALQS LDYALQAFE FRVGAADRG 540
PALSSSALVR VLVLDANDNS PFVLYPLQNG SAPCTELVPR AABEGYLVTK VVAVDGDSGQ 600
NAWLSTYQLK ATEPLPGVW AHNGEVRTAR LLRRERDAKQ RLVLVVKING EPPRSATATL 660
HVLLVDGFSQ FYLLLPFAAP AQQAADLLTV YLVVALASVS SLFLEFVLLF VAVRLCRRSR 720
AASVGRCSVP EGPFPGQMDV VSGTGLSQS YQYEVCLTGG SGTNEFKFLK PIENFVAQG 780
AERVSEANPS PRKSFEST
  
```

Seq ID NO: 289 DNA sequence
 Nucleic Acid Accession #: NM_018674.1
 Coding sequence: 390..2009

1 11 21 31 41 51
 | | | | |

	CGGAGCACAT	GCTGAGCGGA	GCGGCTGGGG	CTGCGCGGCG	TGGCGGAGCA	GCGCTCGCTC	60
	CCTCGCTCAC	TGCTGCTGTC	GCAGGGACAC	ACGCAGGGGC	TGACAGCTGT	GCTGGTGTCTG	120
	ATAAGGGAG	CCACAAGGAG	ACGATCGAGG	AGAGAGACAA	GCGGCAGCAG	AGGCAGCAGC	180
5	GGCAGAGGCA	GCACAGGGC	TGCGGAGCTG	CTGGGAGTGG	GAGTGACTCC	CCCACCTCGG	240
	GCCCCACCC	TGTCCTGTTC	CTCTTCCCGC	TTGCCCTGAG	TTTAGAGAG	CAGCCGCTGC	300
	CACCACTGCC	ACTCGGAGG	GCACCAAGGC	TGCTGGCTAG	GGAGGGACAG	GGCAGGGAGG	360
	CTCTGGCCAG	TCCAGCAGC	CGGGGACAGA	TGCGATCGA	GATTGTGTGC	AAAATCAAAT	420
	TGTCTGAGGA	GGATGCGAAA	CCCAAGGAGA	AGGAGGCAGG	GGATGAGCAG	AGCCTCCTCG	480
10	GGGCTGTTC	CCTTGGAGCA	GCCCCCGGAG	ACCTGGCCAC	CTTTGGCAGC	ACCAGCAGCC	540
	TGCATGAGCT	GGGCCGCGCC	TGTGGGCCAG	GCCCCACGG	ACTGCGCAGA	ACCTGTGTGG	600
	CACCTGGCCCT	ACTCACCTCG	CTGGCTGCCCT	TCCGTACCA	GGCGGCTGGC	CTGGCCCGGG	660
	GCTACTGTAC	CGGCTCTCAC	CTGGTGGCAA	TGGACCCGCG	TGCCCCAGCC	CCAGTGGCGG	720
	GCTTCCCGGC	TGTCACTCTC	TGCAATATCA	ACCGCTTCCG	GCATTCCGCA	CTCAGCGATG	780
15	CGACATCTTT	CCACCTGGCC	AACTGTACAG	GGCTGCCCCC	CAAGAGCCCG	GATGGGCACC	840
	TGCGGCTGCA	GGGCCGCTAC	CCAGAGCCTG	ACATGGTAGA	CATCTCAAC	CGCATGGGCC	900
	ACCACTCGCG	CGACTGCTTT	AAGAGCTGCA	ACTTCAGTGG	GCATCACTGC	TCCGCCAGCA	960
	ACTTCTCTGT	CGCTATATCT	CGCTATGGGA	AGTGTACAC	CTTCAACGCG	GACCCGCGGA	1020
	GCTCGCTGCC	CAGCCGCGCA	GGGGCATGG	GCAGTGGCCT	GGAGATCATG	CTGACATCC	1080
20	AGCAGGAGGA	GTACTCTGCC	ATCTGGAGGG	AGACAAATGA	GACGTGCTTT	GAGGCAGGTA	1140
	TTGGGTGTGA	GATCCACAGC	CAGGAGGAGC	CGCCCTACAT	CCACCAGCTG	GGTTTCGGGG	1200
	TGTCGCCAGG	CTTCCAGACC	TTTGTGTCTT	GCCAGGAACA	GCGGCTGACC	TACCTGCCCC	1260
	AGCCCTGGGG	AGCCGCTGCG	GCAGAGAGTG	AGCTCAGGGA	GCCTGAGCTT	CAGGGCTACT	1320
	CGGCTACAG	TGTGTCTGCC	TGCCGCTGCG	GCTGTGAAAA	GGAGGCCGTG	CTTCAGCGCT	1380
25	GCCACTGCCG	GATGGTGCAC	ATGCCAGGCA	ATGAGACCAT	CTGCCACCA	AATATCTACA	1440
	TGAGTGTGTC	AGACCAACA	CTGGACTCC	TGGGTGGGG	CCCTGAGGGC	CCGTCTTCT	1500
	GCCCCACCC	CTGCAACTCG	ACAGCTATG	GGAAAGAGAT	CTCCATGGTC	AGGATCCCA	1560
	ACAGGGCTTC	AGCCGCTGCG	CTGGGAGGGA	AGTACAACCG	CAACGAGACC	TACATACGGG	1620
	AGAACTTCCT	GGTCTTAGAT	GTCTTCTTTG	AGGCCCTGAC	CTCTGAAGCC	ATGGAGCAGC	1680
30	GAGCAGCCTA	TGGCTGTGCA	GCCCTGTGCG	GAGACCTCGG	GGGACAGATG	GGCCTGTCTA	1740
	TTGGGGCCAG	GATCCACAGC	TTGCTGGAGA	TCTCTGACTA	CATCTATGAG	GTGTCTCTGG	1800
	ATCGACTGAA	GCGGGTATGG	AGGCGTCCCA	AGACCCCTCT	GCGGACTTCC	ACTGGGGGCA	1860
	TCTCACTTTC	GGGCTCTCAG	GAGCTGAAGG	AACAGAGTCC	CTGCCCGAGC	CTGGCCCGAG	1920
	CGGAGGGTGG	GGGGTCTCAG	AGTCTGCTCC	CCAATCAGCA	CCACCCCGAC	GGTCCGCCAG	1980
35	GAGGTCTCTT	TGAAGATTTT	GCCTGTCTAGG	ACGGTGTCTG	GACTGAAAGG	ACCCAGGAGT	2040
	CTGGGACCCC	TCTGGGATC	CCCAGCACAT	TCTCTGCTC	CTGGGAGAGG	CCTGGGGGCG	2100
	GTGCTCACTG	GGAGGGCCAG	GACTCAGTTC	CTGCTCTCAT	CCTCCCTGCG	CCTGATGTCA	2160
	CTGCTTCTTC	ACAAAGTTC	TTCTGTCCA	CACCCCTTAT	CCCCAGGCTG	GTGCCCGGGG	2220
	AGGGCTGGAG	ACCAGGCCAT	GGGCCCTCAC	GGAGAGGAAG	GGAGGAAGG	AGAGGGAGGG	2280
40	GGAGGATAGA	GCCATCCCA	GCCGGGGAGG	GGGAGCCCTC	TGTACATTTC	TAAATATTTC	2340
	GGGAAAGCCG	GATGGGGGGA	GGGATACAG	ATGTAGAAGG	TGGGTAGGGC	TACAGGGGTG	2400
	GGTGATTTAG	GGACAGCCAG	GGTCCAGGCC	CCAATGTCA	CAGGATAGGG	AGAGCCCGAG	2460
	GACTCAGGAG	TGCTGGGCTG	GTCTCTCTTC	CTGCCCTCTC	CCAGGCCAGG	CTCCCTCTTC	2520
	GGCAGGGGGA	GAGGATGGCC	CAGCAGGCCT	GGCCAGCTC	CCAGTTCGCC	CTGCACAGC	2580
45	CCCAACCCCTA	GAGTCCCTTC	TATAGGGAGG	GGGCAGGAGA	CCTTCAGAC	TTCCGGCTGAG	2640
	CTTGGAGGGT	GGGAGGGAG	CCTTCTCAGT	CCTTCTCTCC	TCCAGTCTGA	TTTATATAAG	2700
	TGCTACAGAG						

Seq ID NO: 290 Protein sequence
Protein Accession #: NP_061144.1

50	1	11	21	31	41	51	
	MFIEIVCKIK	FAEEDAKPKE	KEAGDEQSLI	GAFAFGAAPR	DLATPASTST	LHGLGRACGP	60
55	GPHELRRLTL	ALALLTSLAA	FLYQAGLAR	GYLRPHLVA	MDPAAPAPVA	GPPAVTLCNI	120
	NRFRSALSD	ADIFHLANIT	GLPPHORDGH	RAAGLYRPEP	DMVDILNRTG	HQLADMLESC	180
	NFSCEHCAS	NFSVYVTRYG	KCYTFNADER	SSLFSRAGNM	GGGLTMDLID	QEEYLPFWR	240
	ETNETSFEAG	IRVQHSQSE	PPIYHQLGFG	VSPGFQTFVS	QGBRLTYLPE	QFWGNCRABE	300
	ELREPELQGY	SAYSVAACRL	RCEKEAVLQR	CECRMVHMPG	NETICPFNIY	IECADHTLDS	360
60	LGSGPBGFCF	CPTPCNLTRY	GKEISMVRIP	NRGSARYLAR	KYNENETIYR	ENFLVLDPVP	420
	EALTEAMEQ	RAAYGLSALL	GDLGGQGLF	IGASILTLE	ILDYIYEVSW	DRLEKRVNRRP	480
	KPLRLSTGG	ISTLGLQELK	BQSPCFSLGR	ABGGVSSLL	PNEHHPEGPP	GGLFEDFAC	

Seq ID NO: 291 DNA sequence
Nucleic Acid Accession #: Bos sequence
Coding sequence: 62..895

65	1	11	21	31	41	51	
	CACCTGCTCTG	AGAATTGTG	AGCAGCCCTT	AACAGGCTGT	TACTTCCTA	CAACTGACGA	60
70	TATGATCATC	TTAATTFACT	TATTTCTCTT	GCTATGGGAA	GACACTCAAG	GATGGGGATT	120
	CAAGGATGGA	ATTTTTCATA	ACTCCATATG	GCTTGAACBA	GCAGCCGGTG	TGTACCACAG	180
	AGAGCACGCG	TCTGGCAAT	ACAAGCTCAC	CTAAGCAGAA	GCTAAGGCGG	TGTGTGAATT	240
	TGAAGGCGGC	CTCTCGCAA	CTTACAAGCA	GCTAGAGGCA	GCCAGAAAA	TGGATTTC	300
	TGCTGTGCT	GCTGGATGGA	TGGCTAAGGG	CAGAGTTGSA	TACCCCATG	TGAAGCCAGG	360
75	GCCCAACTGT	GGATTGTGAA	AAACTGGCAT	TATGTATTAT	GGAACTCGTC	TCAATAGGAG	420
	TGAAGATGG	GATGCCCTATT	GCTACAACCC	ACACGCAAG	GAGTGTGGTG	GCGTCTTTAC	480
	AGATCCAAAG	CAAAATTTTA	AATCTCCAGG	CTTCCCAAT	GAGTACGAG	ATAACCAAT	540
	CTGCTACTGG	CATATTAGAC	TCAAGTATGG	TCAAGCTATT	CACCTGAGTT	TTTATAGATT	600
80	TGACCTTGAA	GATGACCCAG	GTTGCTTGGC	TGATTATGTT	GAAATATATG	ACAGTTACGA	660
	TGATGTCAT	GGCTTTGTGG	GAGATACCTG	TGGAGATGAG	CTTCCAGATG	ACATCATCAG	720
	TACAGGAAT	GTCATGACCT	TGAAGTTTCT	AAGTATGCT	TCAAGTACAG	CTGGAGGTTT	780
	CCAAATCAAA	TATGTTGCAA	TGATCCCTGT	ATCCAAATCC	AGTCAAGGAA	AAAATACAG	840
	TACTACTTCT	ACGGAATAA	AAAATTTT	AGCTGGAAGA	TTTAGCCACT	TATAAAAAA	900
	AAAAAAGGA	TGATCAAAAC	ACACAGTGT	TATGTTGAA	TCITTTGAA	CTCCTTTGAT	960

CTCACGTGTA TTATTAACAT TTATTTATTA TTTTCTAAA TGTGAAAGCA ATACATAATT 1020
 TAGGGAATAA TGGGAATAT AGGAACCTT AAACGAGAAA ATGAAACCTC TCATAATCCC 1080
 ACTGCATAGA AATAACAAGC GTTAACATT TCATATTTT TCTTTTCAGT CATTTTCTTA 1140
 TTTGTGGTAT ATGTATATAT GTACCTATAT GTATTTCGAT TTGAAATTTT GGAATCCTGC 1200
 TCTATGTACA GTTTGTATT ATACTTTTA AATCTTGAAC TTTATAACA TTTCTCGAAA 1260
 TCATTGATTA TTCTACAAA ACATGATTTT AAACAGCTGT AAAATATTCT ATGATATGAA 1320
 TGTTTTATGC ATTATTTAAG CCTGTCTCTA TTGTGGAAT TTCAGGTGAT TTTCTATAAT 1380
 ATTGTTGCAA TAAATATCCT TGRACACACA AAAAAAAAAA AA

Seq ID NO: 292 Protein sequence
 Protein Accession #: Eos sequence

1 11 21 31 41 51
 MIILYLFL LWDITQWGF KQGIPIHNSIW LERAAGVYHR EARSQKYKLT YAEAKAVCEP 60
 EGGHLATYKQ LEAARKIGFH VCAAGWMARG RVGYPIVKPG PNOGFGKTGI IDYGIKLNRS 120
 ERWDAYCYNP HAKSCGGVFT DPKQIFKSPG FPNEYEDNQI CYWHIRLKYG QRIHLSFLDF 180
 DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELPPDIIS TGNVMTLKFL SDASVTAGGF 240
 QIKYVAMDFV SKSQGKNIS TTSTGNKNFL AGRFSL

Seq ID NO: 293 DNA sequence
 Nucleic Acid Accession #: NM_007115.1
 Coding sequence: 69..902

1 11 21 31 41 51
 GAATTCGCAC TGCTCTGAGA ATTTGTGAGC AGCCCTTAAC AGGCTGTAC TTCCTACAA 60
 CTGACGATAT GATCATCTTA ATTTACTTAT TTCTCTTGCT ATGGGAAGAC ACTCAAGGAT 120
 GGGGATTCAA GGATGGAATT TTTCAATACT CCATATGGCT TGAACGAGCA GCCGGTGTGT 180
 ACCACAGAGA AGCAGGGTCT GGCAATACA AGCTCACCTA CGCAGAAGCT AAGGCGGTGT 240
 GTGATTTGA AGCGCGCCAT CTCGCAACT ACAAGCAGCT AGAGGCAGCC AGAAAAATTG 300
 GATTTCTGT CTGTGCTGCT GGATGGATGG CTAAAGGCGAG AGTTGGATAC CCCATTGTGA 360
 AGCCAGGGCC CAATGATGTA TTTGGAAAAA CTGGCATTAT TGATTATGGA ATCCGCTCTCA 420
 ATAGGAGTGA AAGATGGGAT GCCATTGTCT ACRAACCCACA CGCAAGGAG TGTGTTGGCG 480
 TCTTTACAGA TCCAAAGCGA ATTTTAAAT CTCAGGCTT CCCAATGAG TACGAAGATA 540
 ACCAAATCTG CTACGGCAC ATTAGACTCA AGTATGGTCA GGTATTCAC CTGAGTTTTF 600
 TAGATTTTGA CCTTGAAGAT GACCAGGTT GCTTGGCTGA TTATGTTGAA ATATATGACA 660
 GTTACGATGA TGTCCATGGC TTTGTGGGAA GATACTGIGG AGATGAGCTT CCAGATGACA 720
 TCATCAGTAC AGGAATATGC ATGACCTTGA AGTTTCTAAG TGATGCTTCA GTGACAGCTG 780
 GAGGTTTCCA AATCAATAT GTTGCAATGG ATCCTGTATC CAAATCCAGT CAAGGAAAAA 840
 ATACAGTAC TACTTCTACT GGAATAAAA ACTTTTAGC TGAAGATTAT AGCCACTTAT 900
 AAAAAAAAAA AAGGATGATC AAAACACACA GTGTTTATGT TGGATCTTT TGGAACTCCT 960
 TTGATCTCAC TGTATTATAT AACATTATAT TATTATTTTT CTAAATGTGA AAGAAATACA 1020
 TAAATTAGGG AATATTGGAA AATATAGGAA ACTTTAAAG AGAAAAATGA ACCTCTCATA 1080
 ATCCCACTGC ATAGAAATAT CAAGCGTAA CATTTTCTTA TTTTTCCTT TCAGTCATT 1140
 TTGTATTGT GGTATATGTA TATATGTACC TATATGTATT TGCATTGAA ATTTTGGAA 1200
 CCGCTCTAT GTACAGTTT GTATTATCT TTTTAAATCT TGAACCTTAT GAACATTTC 1260
 TGAATCATT GATTATCTA CAAAAACATG ATTTTAAACA GCTGTAAAAT ATTCTATGAT 1320
 ATGAATGTT TATGCAATAT TTAAGCCGT CTCTATTGTT GGAATTTCAG GTCATTTTCA 1380
 TAAATATGT TGCATTAAT ATCCTTCGGA ATTC

Seq ID NO: 294 Protein sequence
 Protein Accession #: NP_009046.1

1 11 21 31 41 51
 MIILYLFL LWDITQWGF KQGIPIHNSIW LERAAGVYHR EARSQKYKLT YAEAKAVCEP 60
 EGGHLATYKQ LEAARKIGFH VCAAGWMARG RVGYPIVKPG PNXXFGKTGI IDYGIKLNRS 120
 ERWDAYCYNP HAKSCGGVFT DPKRIFKSPG FPNEYEDNQI CYWHIRLKYG QRIHLSFLDF 180
 DLEDDPGCLA DYVEIYDSYD DVHGFVGRYC GDELPPDIIS TGNVMTLKFL SDASVTAGGF 240
 QIKYVAMDFV SKSQGKNIS TTSTGNKNFL AGRFSL

Seq ID NO: 295 DNA sequence
 Nucleic Acid Accession #: NM_001218.2
 Coding sequence: 116..1180

1 11 21 31 41 51
 GTACTGSCCA CGGCAACCCG GCTGCGGCGA CGCGGTCCCG GTGTGCAGCT GGAGAGOGAG 60
 CGGCTACCCG GAGCCCCCGG CACAGCCCGC GCGCGCCCGG CAGGAGCCCG CGAAGATGCC 120
 CGGCGCAGC CTGCAAGCGG CGGCGGTGCT CCGCTGGTG ATCTTAAAGG AACAGCCTTC 180
 CAGCCCGGCC CCAGTGAACG GTTCCAAGTG GACTTATTTT GGTCTGTATG GCGAGAATAG 240
 CTGGTCCAGT AAGTACCGGT CGTGTGGGGG CCGCTGCGAG TCCCCATAG ACCTGACAG 300
 TGACATCCTC CAGTATGAGG CCAGCCTCAC GCGCCCTGAG TTCCAGGCT ACATCTGTC 360
 TGCCACAAGC CAGTTTCTCC TGACCAACAA TGCCCATPCA GTGAAGCTGA ACCTGCCCTC 420
 GGACATGCAC ATTCAGGCGC TCCAGTCTCG CTACAGTGCC AGCAGCTGCG ACCTGCACTG 480
 GGGGAACCCG AATGACCCGC ACGGCTCTGA GCACACCCGC AGCGGACAGC ACTTCGCCGC 540
 CGAGCTGCAC ATTGTCCATT ATAACCTAGA CCTTATCTT GACGCCAGCA CTGCCAGCAA 600
 CAAGCTCAGG GCGCTCGCTG TCCTGGCTGT TCTCATGAG ATGGGCTCTT TCAATCCGT 660
 CTATGACAGG ATCTTCAGTC ACCTTCAACA TGTAAAGTAC AAAGGCCAGG AAGCATTCGT 720
 CCGGGATTC AACATTGAAG AGCTGCTTCC GGAGAGGAGC GCTGAATATT ACCGCTACCG 780
 GGGGTCCCTG ACCACACCCC CTTGCAACCC CACTGTGCTC TGGACAGTTT TCCGAAACCC 840
 CGTGCAAAAT TCCAGGAGC AGCTGCTGCG TTTGGAGACA GCCCTGTACT GCACACACAT 900

5
10
15
20
25
30

```

GGACGACCCCT TCCCCAGAG AAATGATCAA CAACTTCCGG CAGGTCCAGA AGTTCGATGA 960
GAGGCTGGTA TACACCTCCT TCTCCCAAGT GCAAGTCTGT ACTGGGCAG GACTGAGTCT 1020
GGGCATATC CTCTCACTGG CCTGGGCTGG CATCTCTGGC ATCTGTATTG TGGTGGTGGT 1080
GTCCATTGG CTCTTCAGAA GGAAGAGTAT CAAAAAAGGT GATAACAGG GAGTCATTTA 1140
CAAGCCAGCG ACCAAGATGG AGACTGAGGC CCACGCTTGA GGTCCCCGGA GCTCCCGGGC 1200
ACATCCAGGA AGGACCTTGC TTTGGACTCT ACACACTTCG GCTCTCTGGA CACTTGCAGC 1260
ACCTCRAAGT GTTCTCTGTA GCTCAATCTG CAAACATGCC AGGCCTCAGG GATCCTCTGC 1320
TGGGTGCCCT CTTCCTTGG GACCATGGCC ACCCCAGAGC CATCCGATCG ATGGATGGGA 1380
TGCACTCTCA GACCAAGCAG CAGGAATPCA AAGCTGCTTG CTGTAACTGT GTGAGATTGT 1440
GAAGTGGTCT GAATCTTGGG ATCACAACCC AAGCCATGCT GGTGGGCCAT TAATGGTTGG 1500
AAACACTTT CATCCGGGGC TTTGCCAGAG CGTGCTTTCA AGTCTCTGG AAAGTCTGCT 1560
GCTTCTCCAA GCTTTTCAGC AAGAATGTGC ACTCTCTGCT TAGGTTTGG TTGGGAACT 1620
CAACTTCTTT CCTCTGGAGA CCGGGCATCT CCTCTGATT TCCTCTGCT ATGACAAAAC 1680
CTTTAATCTG CACCTTACAA CTCGGGGACA AATGGGGACA GGAAGGATCA AGTTGTAGAG 1740
AGAAAAAGA AACTTACAGG TATACATTGT GATATATTAG GACACTTTC ACAGTCTGCT 1800
CCTCTGGATC ACAGACACTG CACAGACCTT AGGGAATGGC AGGTTCAGAT TCCACTTCTT 1860
GGTGGGGATG AGAAGGAGG GAGAGCTAGA GGGACAAAGA GAATGAGAAG ACATGGATGA 1920
TCTGGGAGAG TCTCTCTTGG GAATCAGAAT TGGAAATCACA TTCTGTTTAT CAAGCCATAA 1980
TGTAAGGACA GAATAATACA ATATTAAGTC CAATCCAC CTCTCTGTCAG TGGAGCAGTT 2040
ATGTTTATA CTCTACAGAT TTTACAATA ATGAGGCTGT TCCTTGAAAA TGTGTGTGTG 2100
CTGTGCTCTG GAGGAGACAT GAGTTCGAG ATGACCCAT CTGCCCTTGA ATCTGGAGGA 2160
AATAGGCAGA AACAAATGA CTGTAGAAT TATTCTCTGT AGGCCAAAT TCATTTCAGC 2220
CACTTCTGCA GGATCCCTAC TGCCAACTCG GAATGGAGAC TTTTATCTAC TTCTCTCTCT 2280
CTGAAGATGT CAATCTGTGG TTAGATCAA ATATATTTC AGCTATAAAA CGCAGAGGTT 2340
ATCTGTGAGG GGGGCTGGCA TCATGTATTT AGGGGCAAGT AATAATGGAA TGCTACTAAG 2400
ATACTCCATA TTCTTCCCG AATCACACAG ACAGTTTCTG ACAGGCGCAA CTCTCCATT 2460
TTCTTCCCGC AGGTGAGAAC CCTGTGGAGA TGAGTCAGTG CCATGACTGA GAAGGAACCG 2520
ACCCCTAGTT GAGAGCACTT TGCAGTTCCC CGAGAAGCTT CTGATTACCA GTCTCAITTT 2580
GACAGCATGA AATGTCTCT TGAAGCATAG CTTTTAAAT ATCTTTTCC TCTACTCTCT 2640
CCCTCTGACT CTAGAATTC TCTCTCTGG AATCGCTGA ACCCAGGAGG CGAGGTTGCT 2700
AGTAAGCCAA GGTATGCCA CTGCACTCTA GCCTGGGTGA CAGAGCGAGA CTCATCTCA 2760
AAAAA AAAA

```

Seq ID NO: 296 Protein sequence
Protein Accession #: NP_001209.1

35
40
45

```

1 11 21 31 41 51
| | | | |
MPRRSLHAAA VLLVLILKEQ PSSPAPVNGS KWTYFQPDGE NWSKKYPSC GGLLQSPIDL 60
HSDLLQYDAS LPLLEFGYIN LSANKQFLLT NMHSHVXLNL PSDMHLQGLQ SRYSATQLHL 120
HWGNPNDRHG SEHTVSGQHF AAKSLHIVHYN SDLYPDASTA SNKSEGLAVL AVLLIEMGSFN 180
PSYDKIPSEL QEVKYKQGA FVPGFNIEEL LPERTAEYR YRSLTTPPC NPTVLMTVFR 240
NPVQISQQL LALETALYCT HMDPSPREM INNFRQVQKF DERLIVYTSFS QVQVCTPAAGL 300
SLGILSLAL ASILGICIV VVSIWLFRRK SIKKGNKGV TYKPAKMET BABA

```

Seq ID NO: 297 DNA sequence
Nucleic Acid Accession #: NM_006632.1
Coding sequence: 377..1582

50
55
60
65
70
75
80

```

1 11 21 31 41 51
| | | | |
ACGGGTCCGC CCACGCGTCC GCCCCGCGGT CCGGTGGGG CCAGAGCGCA GGTGTACCTG 60
CGGGCCGTGC TGGAGCACTT GACGCGCGAG ATCCTGGAGC TGGCTGGCAA CCGGGCCGCT 120
GACAGAAGA CCCGATCAT CTGCGCCAC CTGTAGCTGG CCATTTCGCA CGGCGAGGAG 180
CTTAACAGC TGCTGGGCGA AGTCACCATC GCGCAGGGCG GTGTCTGCTC CAACATTCAG 240
GGGTGCTCT TCCCCAGAA GACCAAGAGC CACCACAAGG CCAAGGGTGA AAACCATTC 300
CTAGGAGAGG AGAAACACAA TGCCACCAA GACAGATTG AGTCCCACAG CAAGGGAGAG 360
CAAGAACGCA CAGATATGC AAGTGGATGA GACACTGATC CCCAGGAAG GTCCAGTTT 420
ATGTTCTGCT CGCTATGAA TAGCCCTCGT CTACATTTC TGCAATTTC CAACGATAGC 480
ACAAAATGTC ATCATGAACA TCACCATGGT AGCCATGGTC AACAGCACAA GCCCTCAATC 540
CCAGCTCAAT GATTCCTCTG AGGTGCTGCC TGTGACTCA TTTGGTGGCC TAAGTAAAGC 600
CCCAAGAGT CTCTCTGCAA AGTCTCAAT ACTTGGGGT CAGTTTGCAA TTTGGGAAA 660
GTGGGGCCCT CCACAGAAC GAAGCAGACT CTGCAGCAT GCTTTATCAG GAATGTACT 720
GGGATGCTTT ACTGCCATCC TCATAGGTGG CTTCATAGT GAAACCCCTG GGTGGCCCTT 780
TGTCTCTAT ATCTTGGAG GTGTTGGCTG TGTCTGCTGC CTCTCTGGT TTGTTGTGAT 840
TTATGATGAC CCTTPTTCT ATCCATGGAT AAGCACTCA GAAAAAGAA ACATCATATC 900
CTCCTTGAAA CACAGGCTG GGTCTCTAA GCAGCCCTT CCAATCAAAG CTATGCTCAG 960
ATCTCTACCC ATTTGGTCCA TATGTTTAG CTGTTTCAGC CATCAATGGT TAGTTAGCAC 1020
AATGGTTGTA TACATACCAA CTATCATCAG CTCTGTGTAC CATGTTAACA TCAGAGACAA 1080
TGGACTTCTA TCTGCCCTTC CTTFATTTGT TGCCGTGGTC ATAGGCATGG TGGGAGGCTA 1140
TCTGGCAGAT TTCTTCTAA CCAAAAAGTT TAGACTCATC ACTGTGAGGA AAATTECCAC 1200
AATTTTAGGA AGTCTCCCC CTTCAGCACT CATGTGTCT CTGCTTACC TCAATTCGCG 1260
CTATATCACA GCAACTGCTT TGCTGACGCT CTCTTGGGA TTAAGCATAT TGTGTGATC 1320
AGGATTTAT ATCAATGTCT TAGATATTGC TOCAAGGAT TCCAGTTTC TCATGGGAGC 1380
ATCAAGAGGA TTTTCAGCA TAGCACCTGT CATGTACCC ACTGTGAGCG GATTTCTTCT 1440
TAGTCAGGAC CTTGAGTTTG GGTGAGGAA TGTCTTCTC TTGCTGTTG CCGTTAACT 1500
GTTAGACTA CTCTCTACC TCAATTTGG AGAAGCAGAT GTCCAAGAT GGGCTAAGA 1560
GAGAAAATC ACTGTTTAT GAAGTTATCC CACCTTGGAT GGAAGATCA TTAGGCACCG 1620
TATTGCTATA AATAGAGGC TTCCGTGATG AAAATACCAG TGAAGAGAT TTTTTCCT 1680
GTGCTCTTT TCAATATGA GATCAGTTCA TTATTTTATT CAGACTTTT TTTGAGAGAA 1740
ATGTAGATG AATAAAAT CAAATAAAT GATAACTAG AAAAAA AAAA

```

Seq ID NO: 298 Protein sequence
Protein Accession #: NP_006623.1

1 11 21 31 41 51
 5 MQVDETLIPR KGPSLCSARY GIALVLHFCN FTTIAQNVIM NITMVAMVNS TSPOSQNLDS 60
 SEVLPVDSFG CLSKAPKSLP AKSSILGGQF AIWEKMGPPQ ERSRLCSIAL SGMLLGCTFA 120
 ILIGGFISST LGWPFVFIYF GGVCVCCLL WFWVIYDDPF SYPWISTSEK EYIISLKKQ 180
 VGSSKQPLPI KAMLRSLPIW SICLGCFSHQ WLVTMVMYI PTYISSVYHV NIRDNGLLSA 240
 LPFIWAVWIG MVGGLADFL LTKKFLITV RKIATILGSL PSSALIVSLP YLNSGYITAT 300
 ALLTLSCGLS TLQSGIYIN VLDIAPRYSS FLMGASRGFS SIAPVIVPTV SGFLLSQDPE 360
 10 FQWRNVFFLL FAVNLLGLLF YLIFGEADVQ ENAKERKLTR L

Seq ID NO: 299 DNA sequence

Nucleic Acid Accession #: NM_003058.1

Coding sequence: 145..1812

1 11 21 31 41 51
 15 GGCCCTGCCC TGAAGGCTGG TCACTTGCAG AGGTAAATCT CCTCTTTGA CTCTGSCCA 60
 GGGTTTGTGC TGAGCTGGCT GCAGCCGCTC TCAGCCTCGC TCGGGGCAAG TCGGGCAGCC 120
 20 TCGGGCCCTC CTGCTTCGAG GATCATGCC ACCACCGTGG ACATGTCCT GGAGCATGGA 180
 GGGGAGTTTC ACTTTTCCA GAAGCAAATG TTTTCTCTCT TGGCTCTGCT CTCGGCTACC 240
 TTGCGGCCCA TCTACGTGGG CATGCTCTTC CTGGGCTTCA CCTGTGACCA CGCTGCGGG 300
 AGCCCCGGAG TGGCCGAGCT GAGTCTGGCG TCGGCTCGGA GTCTGSCAGA GGAAGTGAAC 360
 25 TACAGCGTGC CGGCGCCAGG ACCTGCGGGC GAAGCCTCCC CAAGACAGTG TAGGCGCTAC 420
 GAGGTGAGCT GGAACCAAGG CACCTTTGAC TGGCTGGACC CCTGGCCAG CCTGGACACC 480
 AACAGGAGCC GCGTGGCACT GGGCCCCGCG CGGAGCGGCT GGGGTATCGA GACGCCCTGGC 540
 TCGTCCATTC TGACCTGGGT TAACTGGTA TGTGCCAACT CCTGGATGTT GGACCTATTC 600
 CAGTCATCAG TGAATGTAGG ATTCTTTATT GGCTCTATGA GTATCGGCTA CATAGCAGAC 660
 30 AGGTTTGGCC GTAGCTCTTG CCTCTAACT ACAGTCTCA TAAATGCTGC AGCTGGAGTT 720
 CTCATGCGCA TTCCCCCAAC CTATACGTGG ATGTTAATTT TTCGCTTAAT CCAAGGACTG 780
 GTGAGCAAG CAGGCTGTTT AATAGGCTAC ATCTTGATTA CAGAATTGTG TGGCGCGAGA 840
 TATCGAGAA CAGTGGGGAT TTTTACCA GTTGCCCTATA CAGTTGGGCT CTGGTGCTA 900
 GCTGGGGTGG CTACGCACT TCTCACTGG AGGTGTTGCT AGTTCACAGT TGCTCTGCCC 960
 35 AACTTCTTCT TCTGCTCTA TTAAGTGGTG ATACCTGAGT CTCCAGGTG GCTGATCTCC 1020
 CAGAATAAGA ATGCTGAAGC CATGAGATC ATTAAGCACA TCGCAAGAA AAATGGAAAA 1080
 TCTCTACCCG CCGTCCCTCA GCGCTGAGA CTGGAAGAG AAAGTGGCAA GAAATGAAC 1140
 CCTCATTTCT TGACCTGGT CAGAACTCCT CAGATAAGGA AACATATAT GATATTGATG 1200
 TACAACGGGT TCACGAGCTC TGTGCTCTAC CAGGCGCTCA TCATGCACAT GGGCCTTGCA 1260
 40 GGTGACAATA TCACTCTGGA TTTCTTCTAC TCTGCGCTGG TGAATTGCC AGCTGCCTTC 1320
 ATGATCATCC CACTGAGAG CGCATCGGA GCGCGTACCT CTGGGGTGG ATCAAAATAG 1380
 GTTGCGAGGG CAGCTGCTCT GGCCTCAGTT TTTATACCTG GTGATCTACA ATGGCTAATA 1440
 ATTATTATCT CATGCTTCTG AAGAATGGGG ATCACAATGG CCTATGAGAT AGTCTGCTCG 1500
 GTCATGCTG AGCTGTACCC CACATTCAAT AGGAATCTTG GCTTCCACAT CTGTTCTCTCA 1560
 45 ATGCTGACA TTGGTGGCAT CATCACGCCA TTCTGTTCT ACGGCTCAC TAACATCTGG 1620
 CTGAGCTCC CACTGATGCT TTTCCGCTGA CTGGCTTGG TTGCTGGAGG TCTGGTCTG 1680
 TTGCTTCCAG AAACATAAGG GAAAGCTTG CTTGAGACCA TCGAGGAAGC CGAAATATG 1740
 CAAGACCAA GAATAAATAA AGAAAGATG ATTTACTCC AAGTTCAGAA ACTAGACATT 1800
 CCATTGAAGT AAGAAGAGAG ACCGTGCTG CTGTCATGAC CTAGCTTTGA TGGCAGCAG 1860
 50 ACCAAAGTA GAATCCCTG CACTCATCAC AAAGCCCAT CAACTCAAC AAACCTTACC 1920
 CTGAGCCCTA TCAACTGAG TCTACAGCCA GTGGAGTCTA TTGTACACTG TGGAAAAATA 1980
 CCCATGGGAG CAGATCTGCG CAATTTCTTC CAGCTCACTT TATTCCTAGC ATTCTTAGGA 2040
 CATGGACAT TGGTTTCTG GAGGGTTTTT TTTCCGATCT TTGTATTTT TTAATTTGA 2100
 55 TTCTTTCTT TGCATGCTA GCAACAGAA TACATAGGG AACTGTGGGC TAGGCAANA 2160
 AAATAGAAA AGTGTGAAA ACAGTAAAGT TGGGAGAGGA GCATCTATTT TCTTAAAGAA 2220
 ATAAACACC NAAAACAAA AAAAAA AAAA

Seq ID NO: 300 Protein sequence

Protein Accession #: NP_003049.1

1 11 21 31 41 51
 60 MPTIVDDVLE HGGEHFHFQK QMFLLALLS AIFAPIYVGI VLGFTFDHR CRSPGVARELS 60
 LRCGWSFAHE LMYTVPGPG AGEASPRQR RYEVDWQST FDCVDPLASL DTNRSRLPLG 120
 65 PCRQGVVYET PGSSIVTFEM LVCANSWMLD LPQSSNVGPF FIGSMSIGYI ADRFGRKLCL 180
 LTTVLINAAA GVLMAISPTY TWMLIFRLIQ GLVSKAGWLI GKILITEFVG RRYRRTVGIF 240
 YQVAYTVGLL VLAGVAYALP HNRWLQFTVA LENFFLLLY WCIPSEPRNL ISQNKNAEAM 300
 RIIEKIAKKN GKSLPASLQK LRLEESTGKK LNPFFLDLVR TPQIRKHTMI IMYNWFTSEV 360
 LYQGLIMRMG LAGDNITLDF FYSALVERPA AFNIIITIDR IGRRYFMAAS NMVAGAACLA 420
 70 SVFIPGDLQH LKIIISCIQR MGIIMAYEIV CLVNAELIPT FIEMLGVHIC SSMCDIGGII 480
 TPFLNYRLTN INLELPLMVF GVLGLVAGGL VLLLPETGK ALPETIERBA EMQRPKRNTKE 540
 KMILYQVQL DIPIN

Seq ID NO: 301 DNA sequence

Nucleic Acid Accession #: NM_012206.1

Coding sequence: 52..1131

1 11 21 31 41 51
 75 GTTACCCAGC ATTGTGAGTG ACAGAGCCCTG GATCTGAAGC CTGATCCCAT AATGCATCCT 60
 CAGATGGTCA TCTTAAGCCT CATCTACAT CTGGCAGATT CTGTAGCTGG TCTGTAAAG 120
 80 GTTGTGTGAG AGCAGGTGCC ATCTGTACA CTACCCTGCC ACTACAGTGG AGCTGTACA 180
 TCAATGTGCT GGAATAGAGG CTATGTTCT CTATTCACAT GCCAAATAG CATTGTCTGG 240
 ACCAATGGAA CCCACGTAC CTATCGGAAG GACACAGCT ATAAGCTATT GGGGGACCTT 300
 TCAAGAGGG ATGTCTCTTT GACCATAGAA AATACAGCTG TGTCTGACAG TGGCTATAT 360

5 TGTGCGCGT TTGAGCACC TGCGGTGGTTC AATGACATGA AAATCACCGT ATCATTGGAG 420
 ATTGTGCCAC CCRAGGTAC GACTACTCCA ATGTGCACAA CTGTTCCAAAC CGTCACGACT 480
 GTTCGAAACG GCACCACTGT TCCAAACGACA ACGACTGTTC CAAACGACAA TGTTCCAAACA 540
 ACAATGAGCA TTCCAACGAC AAGCACTGTT CCGACGACAA TGACTGTGTTT AACGACAAACG 600
 AGCGTTCCAA GCACAAACGAG CATTCCAAACA ACAACAAGTG TTCCAGTGAC AACAAACGGTC 660
 TCTACCTTTG TTCTTCCAAT GCCTTTGCCG AGGCAGAAC ATGAACCTAGT AGCCACTTCA 720
 CCATCTTCAC CTCAGCCAGC AGAAACCCAC CCTACGACAC TGCAGGGAGC AATAAGGAGA 780
 GAACCCACCA GCTACCAATT GTACTCTTAC ACAACAGATG GGAATGACAC CGTGACAGAG 840
 TCTTCAGATG GCCTTTGGAA TAACAATCAA ACTCAACTGT TCCTAGAAAC TAGTCTACTG 900
 10 ACGGCCAATA CCACTAAAGG AATCTATGCT GGAGTCTGTA TTCTGTCTTT GTGCTTCTT 960
 GCTCTTTTGG GTGTATCAT TGCCAAAAAG TATTTCTTCA AAAAGGAGGT TCAACAACTA 1020
 AGTGTTCAT TTAGCAGCCT TCAATTAATA GCCTTGCAAA ATGCAGTTGA AAAGGAAGTC 1080
 CAAGCAGAG ACAATATCTA CATTGAGAAT AGTCTTATG CCAACGACTA AGACCCAGTG 1140
 TGTGCTTTTG AGAGTTTACG CCCATGACTG CAGAAGACTG AACAGGTATC AGCAGATCAG 1200
 15 ATGTCTTTTA GAGTCTTACG CAAATTTTCT GTTTCAGTTT CATCTGGCAT TCCAACTATG 1260
 CAGTGATACT GGTGAGATGA ACTCTCCAC TCCAACTGT GTATAGTCAA CCTCATATT 1320
 AATGTAGTGT TAATTTGTTT TGCTAAACT GGCTCAATCC TTCTGATCAT TGCAGAGITT 1380
 TCTCTCAAC ATGAACACTT TAGAATTGTA TGTCTCTTT AGACCCATA AATCTGTAT

Seq ID NO: 302 Protein sequence
 Protein Accession #: NP_036338.1

25 1 11 21 31 41 51
 MHPQVILSL ILHLADSVAG SVKVGSEAGP SVTLECHYSG AVTSMCWNRG SCSLFQONG 60
 IVWTHGTHVT YRKDTRYKLL GDLSRRIVSL TIENAVSDES GVYCCRVEHR GWFNDMKITV 120
 SLEIUPPVVT TPIVITVPT VTIVRTSTTV PTITIVPTT VPTMSIPIPT TIVPTMTVS 180
 TTSVPTTIS IPTTTFVPVT TIVSTFVPPM PLPRQNHFPV ATSPSSFPQA ETHPTTLQGA 240
 30 IRREPTSSPL YSYITDNDT VTESDGLWN NNQTLFLFH SLTANTTKG IYAGVCISVL 300
 VLLALLGVII AKYFIFKEV QQLSVSFSBL QIKALQWAVE KEVQAEADNTY IENSLYATD

Seq ID NO: 303 DNA sequence
 Nucleic Acid Accession #: NM_001044.1
 Coding sequence: 129..1991

35 1 11 21 31 41 51
 ACCGCTCCGG AGCGGGAGGG GAGGCTTCGC GGAACGCTCT CGGCGCCAGG ACTCGCGTGC 60
 40 AAAGCCCGAG CCGCGGGCGG CAGACCAAGA GGAAGAAGC ACAGAAATTC TCAACTCCCA 120
 GTGTGCCCAT GAGTAAGAGC AATGCTCCG TGGGACTCAT GTCTTCCTG GTGGCCCGG 180
 CTAAGGAGCC CAATGCCGTG GGCCTGAAGG AGGTGGAGCT CATCTTGTG AAGGAGCAGA 240
 ACGGAGTGCA GCTCAACAGC TCACCCCTCA CCAACCCGCG GCAGAGCCCC GTGGAGGCC 300
 45 AGGATCGGGA GACCTGGGGC AAGAAGATCG ACTTCTCTCT GTCCGTGATT GGCTTGTCTG 360
 TGGACCTGGC CAACGTCTGG CGGTTCCTCT ACCTGTGCTA CAAAAATGGT GGCAGGTCCT 420
 TCTCTGTCCC CTACCTGCTC TTCAATGGTA TTGCTGGGAT GGCCTTTTC TACATGGAGC 480
 TGGCCCTCGG CCACTTCAAC AGGAAGAGGG CCGCTGGTGT CTGGAAGATC TGCCCCATAC 540
 TGAAAGGTGT GGGCTTCAAG GTCATCTCTA TCACACTGTA TGTGCGCTTC TTCTACAAGC 600
 TCATCATGCG CTGGGCGCTG CACTATCTCT TCTCTCTCT CACCACCGAG CTCCCTCGGA 660
 50 TCCACTGCAA CAATCTCTGG AACAGCCCCA ACTGCTGGGA TGCCCATCCT GGTGACTCCA 720
 GTGGAGACAG CACTGCTCTC AACGACACTT TTGGGACCTAC ACCTGCTGCC GAGTACTTTG 780
 AACGTGGCGT GCTGCACCTC CACGAGAGCC ATGGCATGGA CGACCTGGGG CCTCCGCGGT 840
 GGCAGCTGCT AGCTGCTCAG GTGCTGGTCA TCGTGTGCT CTACTTCAGC CTCTGGAAGG 900
 GCGTGAGAC CTCAGGGAAG GTGGTATGGA TCACAGCCAC CATGCCATAC GTGGTCTCTA 960
 55 CTGCCCCTCT CCTGCTGGGG GTCACTCTCC CTGGAGCCAT AGACGGCATC AGAGCATACC 1020
 TGAGCGTTGA CTCTTACCGG CTCTGCGAGG CGTCTGTTTG GATTGACGCG GCGACCCAGG 1080
 TGTGCTCTCC CCGTGGCGTG GGGTTCGGGG TGCTGATGCG CTCTCCAGC TACAACAAGT 1140
 TCACCAACAA CTGCTACAGG GACGCGATTG TCACCCTTC CATCAACTCC CTGACGAGCT 1200
 TCTCTCTCGG CTCTGCTGTC TTCTCTCTCC TGGGGTACAT GGCACAGAAG CACAGTGTGC 1260
 60 CCATCGGGGA CGTGCCCAAG GACGGGCCAG GCGTGTCTT CATCATCTAC CCGGAAGCCA 1320
 TCGCCACGCT CCCTCTGTTC TCAGCCTGGG CCGTGGTCTT CTTCATCATG CTGCTCACCC 1380
 TGGGTATCGA CAGGCCCATG GGTGGTATGG AGTCAGTGAT CACCGGGCTC ATCGATGAGT 1440
 TCCAGCTGCT GCACAGACAC CGTGAGCTCT TCACGCTCTT CATCGTCTG GCGACCTTCC 1500
 TCTCTCTCCT GTCTCTGCTC ACCAACGGTG GCATCTACGT CTTCACGCTC CTGACCAATT 1560
 65 TTGCAGCGCG CAGGTCCATC CTCTTTGGAG TGCTCATCGA AGCCATCGGA GTGGCTTGGT 1620
 TCTATGTGTT TGGGCAGTTC AGCGACGACA TCCAGCAGAT GACCGGGCAG CGGCCAGCC 1680
 TGTACTGGCG GCTGAGCTGG AAGCTGGTCA GCCCTTCTT TCTCTGTTT GTGGTCTGTT 1740
 TCAGCATGTT GACCTTCAGA CCCCCTACT ACGGAGCTCA CATCTTCCC GACTGGGCCA 1800
 ACGCGCTCGG CTGGGTCTAT GCCACATCTT CCAATGGCAT GGTGCCATC TATGGCGCTT 1860
 70 ACAAGTCTCG CAGCTTGCTT GGGTCTTTC GAGAGAACT GGCCTACGCC ATTGCACCCG 1920
 AGAAGGACCG TGAGCTGGTG GACAGAGGGG AGGTGCGCCA GTTCACGCTC CGCCACTGGC 1980
 TCAGGTGTTA GAGGGAGCAG AGACGAAGAC CCGAGGAAGT CATCTGCAA TGGGAGAGAC 2040
 ACGAACAAAC CAGGAAGATC TAAGTTTCGA GAGAAAGGAG GGCMACTTCT ACTCTTCAAC 2100
 CTCTACTGAA AACACAAACA ACAGAGCAGA AGACTCTCTT CTCTGACTG TTTACACCTT 2160
 75 TCGGTGCGGG GAGCGCACCT CCGCTGTCTT TGTGTTGCTG TAATAACGAC GTAGATCTGT 2220
 GCACGCGTGC CACCCCGGTT GTTGTCCTTG CAGGGCGAGA AAACGTCTAA CTTCATGCTG 2280
 TCTGTGTGAG GCTCCCTCCC TCCCTCTCTT CTGCTCCCGG CTCTGAGGCT GCCCCAGGGG 2340
 CACTGTGTTT TCAGGCGGGG ATCAGATCTT TTGTAGAGCG ACCTGCTTAG AATCCCGGTG 2400
 CTCACAGTAG CTCTCTAGAC CATTACTTTC GCCCATATTA AAAAGCCAAG TGTCTGTCTT 2460
 80 GGTTTAGCTG TGCAGAGGT GAAATGGAGG AAACCACAAA TTCATGCARA GTCTTTTCCC 2520
 GATCGCTGCG TCCAGCAGAA GCGCGTAAAT TGAGCGTICA GTTGACACAT TGCACACACA 2580
 GTCTGTTTCA AGGCATTTGA GGATGGGGGT CCGGTATGTT CTACACAGGA AATTCGTGTT 2640
 ATGTTCTTGG AGCAGAGAGA AATAAACTC CTTGAAACCA GCTCAGGCTA CTGCCACTCA 2700
 GGCAGCTTGT GGTCTCTTGT GTGTAGGGA ACGGCTGAG AGGAGCGTGT CCTATCCCGG 2760
 GACGATGCA GGGCCCCAC AGGAGCGTGT CCTATCCCGG GACGATGCA GGGCCCCAC 2820

5
10
15
20

```

AGGAGCATGT CCTATCCCTG GAGCATGCA GGGCCCCAC AGGAGCGTGT ACTACCCAG 2880
AACGCATGCA GGGCCCCAC AGGAGCGTGT ACTACCCAG GAGCATGCA GGGCCCCAC 2940
TGGAGCGTGT ACTACCCAG GAGCATGCA GGGCCCCAC AGGAGCGTGT CCTATCCCTG 3000
GACCGGACGC ATGCAGGGCC CCCACAGGAG CGTGTACTAC CCCAGGACGC ATGCAGGGCC 3060
CCACAGGAG CGTGTACTAC CCCAGGATGC ATGCAGGGCC CCCACAGGAG CGTGTACTAC 3120
CCACAGGAG ATGCAGGGCC CCCATGCAGG CAGCCTGCAG ACCAACACTC TGCCCTGGCCT 3180
TGAGCGGTGA CCTCCAGGAA GGGACCCAC TGGAAATTTA TTTCTCTCAG GTGCGTGCCA 3240
CATCAATAAC AACAGTTTTT ATGTTTGGA ATGGCTTTTT AAAATCATAT TTACCTGTGA 3300
ATCAAAACAA ATTCAGAAT GCAGTATCCG CGAGCCTGCT TGCTGATATT GCAGTTTTTG 3360
TTTACAAGAA TAATTAGCAA TACTGAGTGA AGGATGTTGG CCRAAAGCTG CTTCCATGG 3420
CACACTGCCC TCTGCCACTG ACAGAAAGT GGATGCCATA GTTTGAATTC ATGCCCTCAG 3480
TGGTGGGCC TGCTACGTG CTGCCGAGG GCAGGGGCCG TGCAGGGCCA GTCATGGCTG 3540
TCCCTGCAG GTGACGTGG CTCCAGGGA CTGGAGTGA ATGCTCGGTG GGAGCGCTCA 3600
GCCTGTGAC TGCCAGGAG CTGCAGTTAG CACAGAGGAT GCCTCCCTCA TTGCCCTCTG 3660
GGGAGGACA CAGAGGACG CTCCCTCATC GCCTCTGCGC CGCTGCAGTC AGCAGAGAGA 3720
GGGCTTCCC CATTCCTTC TGGGAGGGA CACAGAGGAC AGTTTCCCTCA TCGCCTCTG 3780
GTGTTGAAG CACGACAGA GAGCGCTTC CCCATGCGCT TCTGGGAGG GGCTCGGTGT 3840
AGCAACCCAG GTGTGTGCG TGTCGTGTA CCAATCTCTA TTCAGCATCG TGTGGGTCCC 3900
TAAGCAAT AAAAGACATC CACAATGGAA AAAAAAAG GAATTC

```

Seq ID NO: 304 Protein sequence
Protein Accession #: NP_001035.1

25
30
35
40

```

1 11 21 31 41 51
| | | | |
MSKSKCSVGL MSSVVAPEKE PNAVGPKEVE LILVKEQNGV QLTSSILTNF RQSPVEAQR 60
ETWKKIDFL LSVIGFAVDL ANVKRPYLC YKNGGAPLV PYLLFMVIAG MPLFYMELAL 120
GQFMRBGAAG VVKICPILKG VQPTVILISL YVGFFYNVII AMALEYLPSS FTTELPHIHC 180
NNSWNSPNC DAHPGDSBGD SGLNDTFGT TPAARYFERG VLHLHQSHGI DDLPFPWQL 240
TACLVLIVIL LYFLSMKGVK TSKVWVITA TMPYVLTAL LLRGVTLPGA IDIGRAYLSV 300
DFYRLCEASV WIDRAATQCF SIAGVGFLVI AFSSYNKFTN NCYRDAIVTT SINSLSFSS 360
GFVVFSFLGY MAQHSVPIG DVAIDGPGLI FIILPEAIAT LPLSSANAVV FFIMLLTLGI 420
DSAMCGMESV ITGLIDEPQL LHRHRELFIL FIVLATPLLS LFCVINGGIY VFTLLDHFPA 480
GTSLFGVLI EAIGVANFYG VQFSDDIQQ MTGQRPSLYW RLCWKLVSFC FLDFVVVVS 540
VTFPPHYGA YTFPDWANAL GWVIATSSMA MVPIYAAIKG CSLPGSPREK LAYAIAPK 600
RELVDREVR QETLRHMLKV

```

Seq ID NO: 305 DNA sequence
Nucleic Acid Accession #: NM_001216.1
Coding sequence: 43..1422

45
50
55
60
65
70

```

1 11 21 31 41 51
| | | | |
GCCCGTACAC ACCGTGTGCT GGGACACCCC ACAGTCAGCC GCATGGCTCC CCTGTGCCCC 60
AGCCCCCTGGC TCCCTCTGTT GATCCCGGCC CCTGCTCCAG GCCTCACTGT GCAACTGCTG 120
CTGTCACTGC TGCTTCTGAT GCCTGTCCAT CCCACAGAGT TGCCCCGGAT GCAGGAGGAT 180
TCCCTCTGG GAGGAGGCTC TTCTGGGGA GATGACCCAC TGGGCGAGGA GGATCTGCC 240
AGTGAAGAGG ATTCACTCAG AGAGGAGGAT CCACCCGAG AGGAGGATCT ACCTGGAGAG 300
GAGGATCTAC CTGGAGAGGA GGATCTACCT GAAGTTAAGC CTAAATCAGA AGAAGAGGCG 360
TCCCTGAAGT TAGAGGATCT ACCTACTGTT GAGGCTCCTG GAGATCCTCA AGAAGCCAG 420
AATAATCCCC ACAGGACAAA AGAAGGGGAT GACCAGATC ATTGGCGCTA TGGAGGCGAC 480
CGCCCTTGGC CCGGGGTGTC CCGGCTGTC GCGGCGGCT TCCAGTCCCC GGTGGATATC 540
CGCCCTCAGC TCGCCGCTT CTGCCGCGCC CTGCCGCCCT TGGAACTCCT GGGCTTCCAG 600
CTCCCGCGCC TCCAGAACT GCGCCTGCGC AACAAATGGCC ACAGTGTGCA ACTGACCTG 660
CCTCCTGGGC TAGAGATGGC TCTGGGTCCC GGGCGGAGT ACCGGGCTCT GCAGCTGCAT 720
CTGCACCTGG GGGCTGCAGG TGCTCCGGGC TGGAGCACA CTGTGGAGAG CCACCGTTTC 780
CCTCGCGAGA TCCAGTGGT TCACCTCAGC ACCGCTTTG CCAGAGTTGA CGAGGCTCTG 840
GGGCGCCCGG GAGGCTTGGC CGTGTGGCC GCCTTTCTGG AGGAGGGGCC GGAAGAAAC 900
AGTGCCATG AGCAGTTGCT GTCTGCTTG GAAGAAATCG CTGAGGAGG CTCAGAGACT 960
CAGGTCCAG GACTGACAT ATCTGCACTC CTGCCCTCTG ACTTCAGCG CTACTTCCA 1020
TATGAGGGGT CTCTGACTAC ACCGCTCTGT GCCCAGGGT TCATCTGGAC TGTGTTTAA 1080
CAGACAGTGA TGTGAGTGC TAAGCAGCTC CACACCTCT CTGACACCT GTGGGACCT 1140
GGTGACTCTC GGTACAGCT GAACCTCCGA GCGACGAGC CTTTGAATGG GCGAGTGATT 1200
GAGGCTCTCT TCCCTGCTGG AGTGACAGC AGTCTCTGG CTGCTGAGCC AGTCCAGCTG 1260
AATTCTGCTC TGGCTGCTGG TGACATCTTA GCCCTGGTTT TTGGCTCTCT TTTTCTGTC 1320
ACCAGGCTCG CGTTCCTGT GCAGATGAGA AGGCAGCACA GAAGGGGAAC CAAAGGGGGT 1380
GTGAGCTACC GCCCAGAGA GGTAGCCGAG ACTGGAGCCT AGAGGCTGGA TCTTGGAGAA 1440
TGTGAGAGC CAGCCAGAG CATCTGAGG GAGCCCGTA ACTGTCTCTT CCTGCTCAT 1500
ATGCCACTC CTTTAACTG CCAAGAAAT TTTTAAATA AATATTATA AT

```

Seq ID NO: 306 Protein sequence
Protein Accession #: NP_001207.1

75
80

```

1 11 21 31 41 51
| | | | |
MAPLCPSFWL PLLIPAPAPG LTVQLLSLL LIMPVEPQRL PRMQEDSPLG GSSSGEDDPL 60
GEEDLPSEED SPREEDPFG EDLPGEEDLP GEEDLPVVKP KSEBEGSLKL EDLPVEAPG 120
DQEFQNAH RDKEEDDQH WRYGDPFPW RVSPACAGRF QSPVDIRFQL AAFCPALRPL 180
ELLGQPLPPL PELRLRNHG SVQLTLPPGL EMALGPGRY RALQLHLHWG AGRPGSEHT 240
VEGRFPARI HVVRLSTAF RVDRLGRPG GLAVLAAPLE BGPEENSAYE QLLSLREIA 300
BEGSETQVPG LDISALLPSD PSRYFQYEGS LTPPCAQGV INTVFNQTV LSAQLETL 360
DILWPGDSR LQLNFRATP LNRGVLEAF PAGVDSFPA AEPVQLNSCL ARGDILALVF 420
GLLPVATVSA FLVQMRQHR RGIKGVSYR PAEVAETGA

```

Seq ID NO: 307 DNA sequence

Nucleic Acid Accession #: NM_003039.1

Coding sequence: 76..1581

5 1 11 21 31 41 51
 | | | | |
 CTTCTCTCTC CATTCACTGC ACGGTTACT TTGGCTAAAA GGAGGTGAGC GGCACCTCTGC 60
 CCTTCCAGAG CAAGCATGGA GCAACAGGAT CAGAGCATGA AGGAAGGGAG GCTGACGCTT 120
 GTGCTTGCCC TGGCAACCCCT GATAGCTGCC TTTGGGTTCAT CCTTCCAGTA TGGGTACAAC 180
 GTGGCTGCTG TCAACTCCCC AGCACTGCTC ATGCAACAAT TTACAAATGA GACTTACTAT 240
 GGTAGGACCG GTGAATTCAT GGAAGACTTC CCTTGAAGT TGTGTGGTC TGTAAACGGT 300
 TCCATGTTTC CATTGGAGG GTTATCGGA TCCTCTCTGG TCGGCCCTTT GGTGAATAAA 360
 TTTGGCAGAA AAGGGGCTTT GCTGTCAAC AACATATTTT CTATCGTGCC TGGCATCTTA 420
 ATGGGATGCA GCAGAGTCGC CACATCATTT GAGCTTATCA TTATTTCCAG ACTTTTGGTG 480
 GGAATATGTC CAGGTGTATC TTCCAACCTG TCCCATGTG ACTTAGGGGA GCTGGCCCTT 540
 AAAAACCTGC GGGGGGCTCT CGGGTGGTG CCGGAGCTCT TCATCACTGT TGGCATCTTT 600
 GTGGCCAGA TCTTTGGTCT TCGGAATCTC CTTGCAAAAG TAGATGGCTG GCGATCTCTG 660
 CTGGGGCTGA CCGGGGTCCC GCGGGGCTG CAGCTCTCTC TGCTGCCCTT CTTCCCGGAG 720
 AGCCCCAGGT ACTTCTGAT TCAGAAGAAA GACGAAGCGG CCGCAAGAA AGCCCTACAG 780
 ACCCTGCGCG GCTGAGACTC TGTGGACAGG GAGGTGGCCG AGATCCGGCA GGAGGATGAG 840
 GCAGAGAGAG CCGCGGCTT CATCTCCGTG CTGAAGCTGT TCCGATGCG CTOGCTGCGC 900
 TGGCAGCTGC TGTCCATCAT GTCCTCATG GCGCGCCAGC AGCTGTGCGG GGTCAACGCT 960
 ATCTACTACT ACGCGAACCA GATCTACCTG AGCGCGGCGG TGCCGGAGGA GCAGGTGCG 1020
 TACGTGACGG CCGGACCGG GCGCGTGAAC GTGTTCATGA CTTCTGCGCG CGTGTCTGTG 1080
 GTGGAGCTCC TGGTGGAG GCTGCTGCTG CTGCTGGGCT TCTCCATCTG CTTCTAGGCC 1140
 TGCTGCGTGC TCATCTGAGC TCTGGCAGTG CAGGACACAG TGTCTGGAT GCCATACATC 1200
 AGCATCTGCT GCTTCTCTCT CTACGTCATA GGACATGCC TCGGGCCAG TCCATACCC 1260
 GCGCTGCTCA TCACTGAGAT CTTCTGCGAG TCTCTGCGC CATCTGCTT CATGGTGGG 1320
 GGCAGTGTGC ACTGGCTCTC CAATCTACCC GTGGGCTTGA TCTTCCCTTT CATCCAGGAG 1380
 GCGCTGCGCC CGTACAGCTT CATGTCTTTC GCGGTGATCT GCTCTCTCAC CCATCTAC 1440
 ATCTCTCTGA TTGTCCCGGA GACCAAGGCC AAGACGTTCA TAGAGATCAA CCAGATTTTC 1500
 ACCAAGATGA ATAAGGTGTC TGAAGTGTAC CCGGAAAGG AGGAACGTGA AGAGCTTCCA 1560
 CCTGTCACTT CGGAACAGTG ACTCTGAGA GGAAGCCAGT GGAGCTGGTC TGCCAGGGGC 1620
 TTCCCACTTT GGTCTATTTT TCTGACTTCT AGCTGTCTGT GAATATCCAG AAATAAACA 1680
 ACTCTGATGT GGAATGCACT CTTCTCTCTC AGCTCTCCCA CCGCAGTGGG AACTGTGCAA 1740
 AGGGCTGCTT TGTCTCTCTT GAAGCTGGGC TGTCTCTCTC CATGTGGGCC TGTCAACAGA 1800
 CCGAGTCAA TTAACACCT GGTCTCTCAC TTGTCTGTT CAGGCTTGTG GTGGCTCTCT 1860
 GTAACTGGC TCCACCTTGA TGGGTCAACC TTTGTGTGGC TCTGTGTAC ATAACAACA 1920
 CAGTTACTAT AGTGTGAGGA TGGAAAGGAT CAATTTTTC CAGAGAAACT AACTCGTGG 1980
 CCGCAACAGG TCTTCCGGG CCAATGGCAT TTGTAGAG CCAATTCAT CTTCTTACCA 2040
 GATCTCTTTC CAGAAATACC TGTCTAGGAA GGTGTGATGT CAGAAACAAT GACATCCAGA 2100
 AAGCTGAGGA ACAGGTCTCT GTGGAGACAC TGGGTGAGAA TTCTTCATCC AAATATTTT 2160
 GTTAGTGAA AATGAATTG CTTCTGTGTA GTCAATAAAA TGAACCTGAT CACTTTTC

Seq ID NO: 308 Protein sequence

Protein Accession #: NP_003030.1

50 1 11 21 31 41 51
 | | | | |
 MEQQDQSMKE SRLTLVLALA TLIAFGSSP QYGVNVAVN SPALLMQQFY NETYYGRTE 60
 FMEDFPLTLI WSVTVSMFPF GGFIGSLLVG PLVNFPERKG ALLENNIFSI VPAILMGCSR 120
 VATSPFLIII SRLLVGICAG VSENVVPMVL GELAPKHLRG ALGVVQLFI TVGILVAQIF 180
 GLRNLLANVD GWPILLGLTG VFAALQLLLL PFFPESPRYL LIQKDEAAA KXALQTLRGW 240
 DSVDRVARI RQDEDEARKAA GFISVDKLFK MRSLEWQLLS IIVLMGGQQL SGVNAIYYTA 300
 DQIYLSAGVF BEHVQYVTAG TQAVNVMTF CAFVVELLG RRLLLGLPS ICLIAOCVLT 360
 AALALQDTVS WMPYISIVCV ISYVIGHALG PSPIPALLIT EIFLOSRRPS AFNVGSSVHW 420
 LSNFTVGLIF PFIQEHLGEY SFIVFAVICL LTIYIPLIV PETKAKTFIE INQIFKMK 480
 VSEVYPEKRE LKLEPPUTSE Q

Seq ID NO: 309 DNA sequence

Nucleic Acid Accession #: NM_001252.1

Coding sequence: 138..719

65 1 11 21 31 41 51
 | | | | |
 GGCTGGTCCC CTGACAGGTT GAAGCAAGTA GAGGCCAGG AGCCCCGGGA GGGGGCTGCA 60
 GTTTCCTTCC TTCTTCTCG GACGCGCTCC GGCCTCCAT CGCCTCTCT GCGCTAGCGG 120
 AGGTGATGCG CCGCGGATG CCGGAGGAGG GTTCGGGCTG CTCGGTGCAG CCGCGCCCT 180
 ATGGGTGCGT CTTGCGGGCT GCTTGTGCTC CATTTGTCGC GGGCTTGGTG ATCTGCTCTG 240
 TGGTGTGCTT CCAGCGCTTC GCACAGCTC AGCAGCAGCT GCGCTCTGAG TCCTTGGGT 300
 GGGACGTAGC TGAGCTGAGC CTGAATCACA CAGGACCTCA GCAGGACCCC AGGCTATACT 360
 GGCAGGGGGG CCGACGACTG GGCCTCTCTT TCTGCTATGG ACCAGAGCTG GACAAGGGGC 420
 AGCTACGTAT CCATCGTGAT GGCATCTACA TGGTACACAT CCAAGTGAGC CTGGCCATCT 480
 GTCCTCTCAC GACGCGCTCC AGGCACCAAC CACCAACCTT GGCCTGAGGA ATCTGCTCTC 540
 CCGCTCTCCG TAGCATCAGC CTGCTGCTTC TCAGCTTCCA CCAAGGTTGT ACCATTGCTT 600
 CCGAGCGCTT GACGCGCTG GCGCGAGGGG ACACACTCTG CACCAACCTC ACTGGGACAC 660
 TTTGCTCTTC CCGAACACT GATGAGACCT TCTTGGAGT GCAGTGGGTG GCGCCTGAC 720
 CACTGCTGCT GATTAGGCTT TTTTAAATTT TATTTTATTT TATTTAAGTT CAAGAGAAAA 780
 AGTGTACACA CAGGGGCCAC CCGGCGTTTG GGTGGGAGTG TGGTGGGGG TAGTGGTGGC 840
 AGGACAGAG AAGCATTA GCTTTTCTT TCATTTTCTT ATTAATAA

Seq ID NO: 310 Protein sequence

Protein Accession #: NP_001243.1

1 11 21 31 41 51
 | | | | |
 5 MPEEGSGCSV RRRPYGCVLR AALVPLVAGL VICLVVCIQR FAQAQQQLPL ESLGWDVAEL 60
 QLNHTGPOOD PRLYWQSGPA LGRSPLNGPE LDKGQLRIHR DGIYMHVHQV TLAICSSTTA 120
 SRHHPTTLAV GICSPASRSI SLRLSPFQGG CTIASQRLTP LARGDTLCTN LTGTLPLPSRN 180
 TDETFFGVQW VRP

Seq ID NO: 311 DNA sequence

Nucleic Acid Accession #: Eos sequence

Coding sequence: 1..3978

1 11 21 31 41 51
 | | | | |
 15 ATGGTGGGTG AAGGACCCCTA CCTTATCTCA GATCTGGACC AGCGAGGCCG GCGGAGATCC 60
 TTTCAGAGAA GATATGACCC CAGCCTGAAG ACCATGATCC CAGTGCAGCC CTGTGCAGAG 120
 TTAGCACCCA ACCCGGTGGA TGATGCCGGG CTACTCTCCT TCGCCACATT TTCTTGGCTC 180
 ACCCGGTGGA TGGTGAAGAG CTACCGGCAA AGGCTGACCG TAGACACCCT GCCCCCATGG 240
 TCGACATATG ACTCATCTGA CACCAATGCC AAAAGATTTC GAGTCTTTTG GGAAGAGAG 300
 GTAGCAAGGG TGGTCTCTGA GAAGCCCTCT CTGAGCCACG TGGTGTGGAA ATTCAGAGAG 360
 ACACCGGTGT TGATGGACAT CGTGGCCAAAC ATCCTGTGCA TCATCATGGC AGCCATAGGG 420
 CCGACAGTTC TCATTACCA AATCCTCCAG CAGACTGAGA GGACCTCTGG GAAAGTCFGG 480
 GTTGGCATTG GACTGTGCAT AGCCCTTTT GCCACCGAGT TTACCAAGT CTTCTTTTGG 540
 GCCTTGCCT GGGCCATCAA CTACCGCAGC GCCATCCGGT TBAAGGTGGC GCTCTCCACC 600
 TTGGTTTTTG AAAACCTAGT GTCCCTCAGG ACATTGACCC ACATCTCTGT TGGCAGAGTG 660
 25 CTCAATATAC TGATCAAGTA TAGCTATTCT TTGTTGAAG CTGCTTGT TTGCTCTTGG 720
 CCGACCCACA TCCCGATCCT AATGGTCTTT TGTGCGCGGT AGCCCTTTT CATCTGGGG 780
 CCGACAGCTC TCATGCGGAT ATCAGTGTAT GTCATATTCA TACCGTCCA GATGTTTATG 840
 GCCAAGCTCA ATTCAGCTTT CCGAAGGTCA GCATTTTGG TGACAGACAA GCGAGTTCTG 900
 ACNATGAATG AGTTTCTGAC CTGCATCAGG CTGATCAAAA TGTATGCCCT GGAGAAATCT 960
 30 TTTACCAACA TCACCAACAT TATAAGAAGG AGGGAAGAAA AATTACTGGA AAAAGCTGGA 1020
 TTTGTCCAAA GTGGAAACTC TGCCCTGGCC CCCATCTGT CCACCATAGC CATCTGCTG 1080
 ACATTATCCT GCCACATCCT CCTGAGACGC AAACCTACCG CACCCGTGGC ATTTAGTGTG 1140
 ATTGCCATGT TPAATGTAAT GAAGTTTCC ATTGCAATCT TGCCCTCTC CATCAAGCA 1200
 35 ATGGCTGAAG CGAATGTCTC TCTAAGGAGA ATGAGAAAAA TTCTCATAGA TAAAGGCCCC 1260
 CCATCTTACA TCACCAACAC AGAAGACCCA GATACTGTCT TGCTTTAGC AAATGCCACC 1320
 TTGACATGGG AGCATGAAGC CAGCAGGAAA AGTACCCCAA AGAATTTGCA GAACAGAGAA 1380
 AGGCATTAT GCAGAAGACA GAGGTGAGAG GCATACAGTG AGAGGAGTCC ACCAGCCAG 1440
 GGAGCCACTG GCCCAGAGGA GCNAAAGTAC AGCCTCAAT CCGTCTCTGA CAGCATAAGC 1500
 40 TTTGTGGTGA GAAAGTTATG TCGTTATCCC GAAGCCGAGC TOCTGGCTTG GAGGTGGCCA 1560
 GCRGTGTTTG TTGGGGAAT CATCAGAGGA TACAGGCCCT ATGGATTTTC TGCTAAAGAC 1620
 AAGGATGAAT CTAGAAGGCT TCTTACTTGG CCCCAGAGAG TGGATAGGAC TCAAGGGGCA 1680
 CCGAAATACC TGGGGAGAT CTGGGAATA TGTGGGAATG TGGGAAGTGG AAGAGCTTCC 1740
 CTCCTTGCG CPTCTCTAGG ACAGATGCAG CTGCAAGAA GGGTGGTGGC AGTCAATGGA 1800
 45 ACTTTGGCCT AGTTTTCACA GCAGGCATGG ATCTTTCTG GAAATGTGAG AGAAACATA 1860
 CTCTTTGGAG AAAAGTATGA TCACCAAGG TATCAGCACA CAGTCCGCGT CTGTGGCCTC 1920
 CAGAGAGACC TGAGCAACCT CCCCTATGGA GACCTGACTG AGATTGGGGA GCGGGGCCCTC 1980
 AACCTCTCTG GCGGAGAGAG GCAGAGGATT AGCCTGGGCC GCGCTGTCTA CTCGAGCCTG 2040
 CAGCTCTACC TGCTGGAGGA CCCCCTGTGG GCGCTGGAGC CCGACGTGGG GAAGCACCTC 2100
 50 TTTGAGGAGT GCATTAGAA GACGCTCAGG GGAAGACAG TGGTCCYGGT GACCCACCAG 2160
 CTACAGTTCT TAGAGTCTTG TGATGAAGTT ATTTTATTAG AAGATGGAGA GATTGTGAA 2220
 AAGGGAACCC ACAGAGAGTT AATGGAGGAG AGAGGGCGCT ATGCAAAACT GATTCACAA 2280
 CTGCGAGGAT TGCAAGTTCA GGATCCTGAA CACCTTTACA ATGCAGCAAT GGTGGAAGCC 2340
 TTCAAGGAGA GCGCTGCTGA GAGAGAGGAA GATGCTGATA TAATCGGTA CTTCTTTCT 2400
 55 CTCTTCACTG TGTCTCTCTT CTTCTGATG ATTGGCAGCG CTGCTTCTAG CAACTGGTGG 2460
 CTGGGTCTCT GGTGAGACAA GGGCTCAGG ATGACCTGTG GGGCCAGGG CAACAGGACC 2520
 ATGTGTGAGG TCGGCGCGGT GCTGCGAGAC ATCGGTGAGC ATGTGTACCA GTGGGTGTAC 2580
 ACTGCAAGCA TGGTGTTCAT GCTGGTGTGT GCGGTACCA AAGGCTTCT CTTACCAAG 2640
 ACCACACTGA TGGCATCTCT CTTCTGATG GACACGGTGT TGTATAGAT CTTAAGAGC 2700
 60 CCAATGAGTT TCCTTGACAC GACTCCACT GGCAGGCTAA TGAACGTTT TTCCAAGGAT 2760
 ATGACAGAGC TGATGTGAG GCTGCTTCT CACGAGAGA ACTTCTGCA GCAATTTTTT 2820
 ATGGTGGTGT TTATCTCTGT GATCTTGGCT GCTGTGTTTC CTGCTGTCTT TTAGTCTGTG 2880
 GCCAGCCTTG CTGTAGGCTT CTTCATTCTG TTAAGCATTT TCCACAGAG AGTCCAGGAG 2940
 CTCAGAGAGG TGAGAGATGT CAGCGGTCA CCGTGTGTA CCGACATCAC CTCTCTCATG 3000
 65 CAGGCGCTGG GCATCAATCA CGCCTATGGC AAGAAGGAGA GCTGCATCAC CTATCTCTCA 3060
 TCCAAGGCC TGTCAATGTC ATACATCATC CAGCTGAGCG GACTGCTCCA AGTGTGTGTG 3120
 CGAAGCGGAA CAGAGAGGCA AGCCAAATTC ACCTCCGTGG AGCTGCTCAG GGAATACATT 3180
 TGGACCTGTG TTCTCTGAATG CACTCATCCC CTCAGAGTGG GGACCTGTCC CAAGGACTGG 3240
 CCCAGCTGTG GGGAGATCAC CTTGAGAGAC TATCAGATGA GATACAGAGA CAACACCTCC 3300
 70 CTGTGTTCTG ACAGCTGAA CTTGAACATA CAAAGTGGGC AGACAGTCCG GATTGTTGGA 3360
 AGAACAGGTT CCGGAAGTTC ATCGTTAGGA ATGGCTTTGT TCGTCTGTGT GAGCCAGGCC 3420
 AGTGGCACAA TCTTTATTGA TGAGGTGGAT ATCTGCATTC TCAGCTTGA AGACCTCAGA 3480
 ACCAAGCTGA CTGTGATCCC ACAGGATCCT GTCTGTGTTG TAGGTACAGT AAGGTACAA 3540
 75 TTGGATCCTT TTGAGAGTCA CACCGATGAG ATGCTCTGCG AGGTCTGGA GAGAACATTC 3600
 ATGAGAGACA CAATAATGAA ACTCCAGAA AAATTACAGG CAGAACTCAC AGAAAATGGA 3660
 GAAAACCTCT CAGTAGGGGA ACCTCAGCTG CTTTGTGTGG CCGAGCTCT TCTCGTAT 3720
 TCAAGATCA TTCTCTCTGA TGAGGCCACC GCCTCTATGG ACTCCAGAC TGACACCTCT 3780
 GTTCAGAACT CCAATCAAGA TGCTCTCAG GGTGCACTG TGTGACCAT CGCCACCGC 3840
 CTCACACAG TTCTCAACTG CGATCACGTC CTGTTATGG AAAATGGGAA GGTGATTGAG 3900
 80 TTTGACAAGC CTGAAGTCTT TGACAGAGAG CCAGATTCTG CATTTGCGAT GTTACTAGCA 3960
 GCAGAGTCA GATTGTAG

Seq ID NO: 312 Protein sequence

Protein Accession #: Eos sequence

1 11 21 31 41 51
 5 MVGEGEYLIS DLDQGRRRS FAERYDPSLK TMIFVRFCAR LAPNPDVDDAG LLSFATFSML 60
 TPVMVKGYSR RLTVDTLPL STYDSSDTNA KRFRVLWDEE VARVOPEKAS LSHVVMKFOR 120
 TRVLMDIVAN ILCIIMAAIG PTVLIHQILQ QTERTSGKVN VGIGLCIALF ATEFTKVFFW 180
 ALAWAINYRT AIRKVALST LVFENLVSPK TLTHISVGEV LNILSSDSYS LPFAALPCPL 240
 PATIPILMVF CAAYAPFILG PTALIGISVY VIFIPVQMFH AKLNSAFRRS AILVTDKRVQ 300
 TMNEPLTCIR LIRMYAWEKS FTNTIQDIRR RERKLEKAG PVQSGNSALA PIVSTIAIVL 360
 10 TLSCHILLRR KLTAPVAPSV IAMENVMKPS IATLPFSIKA MAEAMVSLRR MKKILIDKSP 420
 PSYITQPEDP DTVILLANAT LTWZHEASRK STPKKLQNK RHLCKQRSE AYSESRPPAK 480
 GATGPESQSD SLKSVLHSIS FVVRKLCRYP EAQLLAWRHF AVFVGRIIRG YRPHGFSKAD 540
 KDESRLILTW PQEVDRTORA AKYLKILGI CGNVGSGKSS LLAALLGQMQ LQKGVVAVNG 600
 TLAYVSQQAQ IPHGNVRENI LFGKDYDHQR YQHTVRVCGI QKDLNMLPYG DLTEIGERGL 660
 15 NLGGGQRQRI SLARAVYSDR QLYLLDDFLS AVDAEVGKEV FECEIKKTLR GKTVVVLVTHQ 720
 LQFLESCDVP ILLEDGEICE KGTHELMES RGRYAKLIEN LRGLQFKDPE HLYNAMVEA 780
 FKESPAEREE DAGIYGLYS LFTVFLFLM IGSAAFSNWW LGWLWDKGRS MTCGPQGNRT 840
 MCEVGAULVD IQGVVYQWYV TASHVFMVLF GVTKGPFVTK TILMASSSLH DTVFDKILKS 900
 PMSFFDTTPT GRIMNRFSD MDELDRLEPF HAENFLQQPF MVVFILVILA AVFPAVLIVV 960
 20 ASLAVGEFLL LRIFHRGQVE LKKVENVRS PWFTHITSM QGLGIIHAYG KKBSCITYTS 1020
 SKGLSLSYII QLSGLLQVCV RTGTETQAKP TSVELLREYI STCVPECTHP LKVGTCPKDW 1080
 PSCGKITFRD YQMYRLDNTF LVLDLNLINI QSGQTVBIVG RTGSGKSSLG MALFRLVEFA 1140
 SCITFIDEVD ICILSLDLR TKLTVIPQDP VLFVGTIVRYN LDFFESHTDE MLWQVLEKTF 1200
 25 MRDTIMKLPE KIQAEVTENO ENFSVGRQL LCVARALLRN SKIILLDEAT ASMDSKTDTL 1260
 VQNTIKDAFK GCTVLTIAHR LNTVLNCDHV LVMENGVIE FDKPEVLAEK PDSAPAMLLA 1320
 AEVLL

Seq ID NO: 313 DNA sequence
 Nucleic Acid Accession #: Z31560
 Coding sequence: 1-966

1 11 21 31 41 51
 30 CACAGCGCCC GCATGTACAA CATGATGGAG ACGGAGCTGA AGCCGCCGGG CCGCAGCAA 60
 ACTTCGGGGG GCGGGGGGG CAACATCCACC GCGGGGGGG CCGGGGGCAA CCGAATAAAC 120
 35 AGCCCGGACC GGTCAAGCG GCCCATGAAT GCCTTCATGG TGTGGTCCCG CCGGCAGCGG 180
 CGCAAGATGG CCGCAGAGAA CCCCAGATG CACAACCTGG AGATCAGCAA GCGCCCTGGC 240
 GCGGATGGG AACTTTTGTG GGAGACGGAG AAGCGGCGT TCATCGACGA GGCTAAGCGG 300
 CTGCGAGCGC TGACATGAA GGAGCACCG GATTATAAAT ACCGGCCCGG GCGGAAACCC 360
 40 AAGACGCTCA TGAAGAAGGA TAAGTACAG CTGCCCGGG GGTCTCTGGC CCGCGGGCG 420
 AATAGCATGG CGAGCGGGGT CGGGGTGGG GCGGGCCCTG GCGCGGGCTT GAACACGCGC 480
 ATGACAGGTT ACCGCGCAT GAAACGCTGG AGCAACGCGA GCTACAGCAT GATGAGGAC 540
 CAGCTGGGT ACCGCGACA CCGGGGCTC AATGCGCACG GCGCAGCGCA GATGAGGCC 600
 ATGACCGCT ACAGCGTGG CGCCCTGCG TACAACCTCA TGACTAGCTC GCAGACCTAC 660
 45 ATGAACGGCT CGCCACCTA CAGCATGTCC TACTGCGAG AGGGCACCCC TGGCATGGCT 720
 CTGGCTCCA TGGGTTCGGT GGTCAAGTCC GAGGCCAGCT CCAGCCCCCT TGTGGTTACC 780
 TCTTCTCCC ACTCCAGGGC GCGCTGCCAG GCGGGGAGCC TCCGGGACAT GATCAGCATG 840
 TATCTCCCG GCGCCCGGT GCGGAACCC GCGGCCCCA GCAGACTTCA CATGTCCAG 900
 CACTACCAGA GCGCCCGGT GCGCGCACG GCCATTAACG GCACACTGCC CCTCTCACAC 960
 50 ATGTGAGGGC CGAGCAGCGA ACTGGAGGGG GGAGAAATTT TCAAGAAAA ACCAGGAAA 1020
 TGGAGGGGT GCAAAAGAGG AGAGTAAGAA ACAGCATGGA GAAACCCGG TACGCTCAA 1080
 AAAAA

Seq ID NO: 314 Protein sequence
 Protein Accession #: CAA83435

1 11 21 31 41 51
 55 HSARMYNDME TELKPFPGPQ TSOGGGGNST AAAAGGNQKM SPDRVKRFMN AFMVNSRGQR 60
 60 RKMAQENPKM HNSEISKRLG AEWKLLSEYE KRPFIDEAKR LRALMKHEHP DYKVRPREKT 120
 KTLMKKDKYT LFGGLLAPGG NSMASGVGVG AGLGAGVNR MDSYAHMNGW SNGSYSMNQD 180
 QLGYPPQHPGL NARGAQMOP MHRVDVSLQ YNSMTSSQTY MNGSPTYSMS YSQGTPGMA 240
 LGSMSGVVKS EASSSPFVVT SSSHSRAPCQ AGDLRDMISM YLPGAEPVPEP AAPSRLEMSQ 300
 HYQSGVPVPGT AINGTLPLSH M

Seq ID NO: 315 DNA sequence
 Nucleic Acid Accession #: U91618
 Coding sequence: 29..541

1 11 21 31 41 51
 70 CGGACTTGGC TTGTTAGAAG GCTGAAAGAT GATGGCAGGA ATGAAATCC AGCTGTATG 60
 CAGCTACTFC CTGGCTTCA GCTCCTGGAG TCTGTGCTCA GATTGAGAAG AGGAATGAA 120
 AGCATTAGAA CGAGATTCTT TGACCAATAT GCATACATCA AAGATTAGTA AAGCACATGT 180
 75 TCCCTCTTGG AAGATGACTC TGCTAAATGT TTGCAGTCTT GTAAATAAT TGAACAGCCC 240
 AGCTGAGGAA ACAGAGAGA TTCATGAAGA GGAGCTTGTI GCAAGAAGGA AACTTCTTAC 300
 TGCTTTAGAT GCTTTTAGCT TGGAGCAAT GTTGACAATA TACCAGCTCC ACAAATCTG 360
 TCACAGCAGG GCTTTTCAAC ACTGGGAGTT AATCCAGGAA GATATTCTPT ATACTGGAAT 420
 TGACAAAAAT GGAAAGGAAG AAGTCATAAA GAGAAAAATT CCTATATTC TGAACAGGCA 480
 80 GCTGTATGAG AATAAACCCA GAAGACCTCA CATACTCAA AGAGATTCCT ACTTACTG 540
 AGAGAATAAA TCATTATATT ACATGTGATT GTGATTATC ATCCCTTAAT TAAATATCAA 600
 ATTATATTG TGAGAAATG TGACAAACAC ACTTATCTGT CTCTTCTACA ATTGTGGTTT 660
 ATTGAATGTG TTTTCTGCA CTAATAGAAA TTAGACTAAG TGTTTTCAA TAAATCTAAA 720
 TCTTCAAAAA AAAAAAAA AAATGGGGCC GCAATT

Seq ID NO: 316 Protein sequence
Protein Accession #: AAB50564

```

5      1      11      21      31      41      51
      |      |      |      |      |      |
MMAGMKIQLV CMLLLAFSSW SLCSDSSEEM KALEADFLTN MHTSKISKAH VPSWKMTLLN 60
VCSLVNMLNS PAETGTGVHE EELVARRKLP TALDGFSLFA MLTIYQLRHI CHSRAPQHWE 120
LIQEDILDG NDKNKKESEVI KRKIPYILKR QLYENKPRRP YILKRDSYYY

```

Seq ID NO: 317 DNA sequence
Nucleic Acid Accession #: NM_006536.2
Coding sequence: 109..2940

```

15      1      11      21      31      41      51
      |      |      |      |      |      |
ACCTAAACCT TTGCAAGTTC AGGAAGAAAC CATCTGCATC CATATTGAAA ACCTGACACA 60
ATGATATGCG CAGGCTCAGT GTGAGTGAAC TGGAGGCTTC TCTACAACAT GACCCAAAGG 120
AGCATTTGCG GTCCATTTTG CAACCTGAAG TTTGTGACTC TCCTGGTTGC CTTAAGTTCA 180
GAACTCCCAT TCCTGGGAGC TGGAGTACAG CTTCAAGACA ATGGGTATAA TGGATTGCTC 240
ATTGCAATTA ATCCTCAGT ACCTGAGAAT CAGAACCTCA TCTCAAACAT TAAGGAAATG 300
ATAACTGAAG CTTCAATTTA CCTATTTAAT GCTACCAAGA GAAGAGTATT TTTCAGAAAT 360
ATAAAGATT TAAATACCTGC CACATGGAAA GCTAATAATA ACAGCAAAAT AAAACAAGAA 420
TCATATGAAA AGGCAAAATGT CATAGTGAAT GACTGGTATG GGGCACATGG AGATGATCCA 480
TACACCTTAC AATACAGAGG GTGTGGAAAA GAGGGAAAT ACATTATTT CACACCTAAT 540
TTCTTACTGA ATGATAACTT AACAGCTGGC TACGGATCAC GAGGCGGAGT GTTTGTCCAT 600
GAATGGGCCC ACCTCCGTTG GGGTGTGTTT GATGATATA ACAATGACAA ACCTTCTTAC 660
ATAAATGGGC AAAATCAAT TAAAGTGACA AGGTGTTTCT CTGACATCAC AGGCATTTTT 720
GTGTGTGAAA AAGGTCCTTG CCCCAGAA AACTGTATTA TTAGTAAGCT TTTTAAAGAA 780
GGATGCACCT TTATCTACAA TAGCACCCAA AATGCAACTG CATCAATAAT GTTCATGCAA 840
AGTTTATCTT CTGTGGTTGA ATTTTGTAAAT GCRAGTACCC ACAACCAAGA AGCACCAAAC 900
CTACAGAACG AGATGTGCGAG CCTCAGAAGT GCATGGGATG TAATCACAGA CTCTGCTGAC 960
TTTCACCACA GCTTTCCTCAT GAATGGGACT GAGCTTCCAC CTCCTCCCAC ATTCCTGCTT 1020
GTACAGCTCG GTGACAAAGT GGTCTGTGTTA GTGCTGGATG TGTCCAGCAA GATGGCAGAG 1080
GCTGACAGAG TCCTCAACT ACAACAAGCC GCAGAAATTT ATTTGATGCA GATTGTTGAA 1140
ATTCTACTCT TCTGTGGCAT TGCCAGTTTC GACAGCAAA GAGGATACAG AGCCGAGCTA 1200
CACCAAAATA ACAGCAATGA TGATCGAAAG TTGCTGGTTT CATATCTGCC CACCACTGTA 1260
TCAGCTAATA CAGACATCAG CATTTGTTC A GGGCTTAAGA AAGGATTTGA GTTGGTTGAA 1320
AAACTGAAAT GAAAGCTTFA TGGCTCTGTG ATGATATTAG TGACCGAGCG AGATGATAAG 1380
CTTCTTGCCA ATTGCTTACC CACTGTGCTC AGCAGTGGTT CAACAATTCA CTCCTATGCC 1440
CTGGGTTCAT CTGCGGCCCC AAATCTGGAG GAATTATCAC GTCTTACAGG AGGTTTAAAG 1500
TTCTTTGTTT CAGATATATC AAATCTCAAT AGCATGATTG ATGCTTTCAG TAGAATTTCC 1560
TCTGGAAGTG GAGCAATTTT CCAGCAACAT ATTCTAGCTT AAGTATCAGG TGAAATATGC 1620
AAACCTCACC ATCAATTGAA AAACACAGTG ACTGTGGATA ATACTGTGGG CAACGACACT 1680
ATGTTTCTAG TTACGTGGCA GGCCAGTGGT CCTCTGAGA TTATATTATT TGATCCTGAT 1740
GAGCGAAAT ACTACACAAA TAATTTTATC ACCAATCTAA CTTTCCGAC AGCTAGTCTT 1800
TGGATTCCAG GAACAGCTAA GCCTGGGCAC TGGACTTACA CCTGGAACAA TACCCATCAT 1860
TCTCTGCAAG CCCTGAAAGT GACAGTGACC TCTCGGCCCT CCAACTCAGC TGTGCCCCCA 1920
GCCACGTGGG AAGCCTTTGT GGAAGAGAGC AGCCTCCATT TTCTCATCC TGTGATGATT 1980
TATGCCAATG TGAACAGGG ATTTTATCCC ATTCTTAATG CCACTGTCCAC TGCCACAGTT 2040
GAGCAGAGA CTGAGGCTT TGTACGCTG AGACTCCTTG ATGATGGAGC AGGTGCTGAT 2100
GTTATAAAAA ATGATGGAAT TTACTCGAGG TATTTTCTCT CCTTGTCTGC AAATGTTAGA 2160
TATAGCTTGA AAGTGCATGT CAATCACTCT CCCAGCATAA GCACCCGAGC CCACTCTATT 2220
CCAGGGAGTC ATGCTATGTA TGTACCAAGT TACACAGCAA ACGGTAATAT TCAGATGAAT 2280
GCTCCAAGGA AATCAGTAGG CAGAATGAG GAGGAGCGAA AGTGGGGCTT TAGCCGAGTC 2340
AGCTCAGGAG CTCTCTTTC AGTCTGGGA GTTCCAGCTG GCCCCACCC TGTGTGTTT 2400
CCACCATGCA AAATATTGGA CTTGGAAGCT GTAAAGTAG AAGGGAATT GACCCATCT 2460
TGGACAGCAC CTGAGAAGA CTTTGTATCAG GGCCAGGCTA CAAGCTATGA AATAAGATG 2520
AGTAAAGTC TACAGATAT CCAAGATGAC TTTAACAATG CTATTTTAGT AATACATCA 2580
AAGCGAAATC CTCAGCAAGC TGGCATCAGG GAGATATTTA CGTTCTCACC CCAAGATTCC 2640
AOGAATGGAC CTGAACATCA GCCAATGGA GAAACACATG AAAGCCACAG AATTATGTT 2700
GCARTACGAG CAATGATAG GAATCCTTA CAGTCTGCTG TATCTAACAT TGCCGAGCG 2760
CCTCTGTTTA TTCCGCCCAA TTCTGATCCT GTACCTGCCA GAGATTATCT TATATTGAA 2820
GGAGTTTAA CAGCAATGGG TTTGATAGGA ATCATTGCC TTATTATAGT TGTGACAT 2880
CATACTTAA GCAGAAAAA GAGAGCAGAC AAGAAAGAGA ATGGAACAAA ATTATTATA 2940
ATAAATATCC AAGGTGTCTT CCTCTTAGA TATAAGACCC ATGGCCTTCG ACTACAAAA 3000
CATACTAACA AAGTCAATT AACATCAAAA CTGTATTAAA ATGCATTGAG TTTTGTACA 3060
ATACAGATAA GATTTTACA TGGTAGATCA ACAATCTTT TTGGGGTAG ATTAGAAAAC 3120
CCTTACACTT TGGCTATGAA CAATAATAA AAATTATTCT TTAAAGTAAT GTCTTTAAAG 3180
GCAAGGGGAA GGGTAAAGTC GGACCAAGT CAAGGAAAGT TTGTTTATT GAGGTGGAAA 3240
AATAGCCCCA AGCAGAGAAA AGGAGGGTAG GTCTGCATTA TAAGTGTCTG TGTGAAGCAA 3300
TCATTTAGTT ACTTTGATTA ATTTTCTT TCTCCTTATC TGTGCAATAC AGGTGTCTG 3360
TTTACATGAA GATCATGCTA TATTTTATAT ATGTAGCCCC TAATGCAAG CTCCTTACCT 3420
CTGTCTATT TTTTATATAT ATTTACAGAT ACATCTCCCT GCTAATGCTC AGAGATCTTT 3480
TTTCATGTA AGAGGTAACC TTTAACAATA TGGGTATTAC CTTTGTCTCT TCATACCGGT 3540
TTTATGACAA AGGTCTATTG AATTATTG TGTGTAAGT TCTACTCCA TCAAGCAGC 3600
TTTCTAAGTT TATGCTTGT GGTATTATG GAATGATAGT TATAGCCCN TATAATGCTT 3660
TACCTAGGAA A

```

Seq ID NO: 318 Protein sequence
Protein Accession #: NP_006527.1

```

1      11      21      31      41      51
      |      |      |      |      |
MTQRSIAGPI CNLKFVILLV ALSELPLFG AGVQLQDNGY NGLLIAINPQ VPEQNQLISN 60

```

IKEMITEASF YLFPNATKRRV EFRNIKILIP ATWKANNNSK IKQESYEKAN VIVTDWYGAR 120
 GDDPYTLQVR GCGKKGKXIH FTFNPLLDN LTAGYGSRRR VVVEHWAHLR WGVFDEYNND 180
 KPFYINGQNO IKVTRCSSDI TGIFVCEKGP CPOENCIIK LFKSGCTFIY NSTQNATASI 240
 MFMSLSVV EFCNASTBNQ EAPNLQNMOC SLRSANDVIT DSADFHHSFP MNGTELPPPP 300
 5 TFSLVQAGDK VVCLVLDVSS KMAEADRLLO LQQAAREFYLQ QIVEIHTFVG IASFDKSGEI 360
 RAQLHQINSN DDRKLLVSYL FTTVSAKTDI SICSGLKXGF EVVEKLNGKA YGSVMILVTS 420
 GDDKLLGNCL PTVLSSGSTI HSIALGSSAA FNLEELSRIT GGLKPFVVDI SNSNSMIDAF 480
 SRISSGTGDI FQHQHLEST GENVKPHHQL KNTVTVDNTV GNDTMTPLVW QASGPPEIIL 540
 10 FDDPGRKYIT NNFITNLTPR TASLWIPGTA KPGHWTYTLN NTHSLQALK VTTSTRASNS 600
 AVPPATVEAF VERDSLHFPF PVMYANVKQ GFYPILNATV TATVEPETGD PVTLRLLDDG 660
 AGADVIKNDG IYGRYFSPFA ANGRYSKVE VNHSPSISTP AHSIFGSHAM YVPGYTANGN 720
 IQMNAPRKSV GRNEERKMG FSRVSSGGSF SVLGVPAGPH PDVFPCKII DLEAVKVEBE 780
 LTLSTWAPGE DFDQQAQTSY EIRMSKSLQN IQDDFNAIL VNTSKRNPQO AGIREIPTFS 840
 15 PQLSTNGPEH QPNGETHSH RIYVAIRAMD RNSLQSAVSN LAQAPLFIPI NSDFVPARDY 900
 LILKGVLTAM GLIGIICLI VVTHHTLSRK KRADKKENG T KLL

Seq ID NO: 319 DNA sequence

Nucleic Acid Accession #: NM_000228.1

Coding sequence: 82..3600

20 1 11 21 31 41 51
 GCTTTCAGGC GATCTGGAGA AAGAACGGCA GAACACACAG CAAGGAAAGG TCCTTTCTGG 60
 GGATCAGCCC ATTAGGCTGAA GATGAGACCA TTCTTCTCT TGTGTTTTC CCTGCCTGGC 120
 25 CTCTGTCATG CCAACAACAG CTGCTCCCGT GGGGCCCTGT ATCCACCTGT TGGGGACCTG 180
 CTTGTTGGGA GGACCCGGTT TCTCCGAGCT TCATCTACCT GTGAGCTGAC CAAGCCTGAG 240
 AACTACTGCA CCAGATATGG CGAGTGGCAG ATGAAATGCT GCAAGTGTGA CTCAGGCGAG 300
 CTTCAAACT ACTACATGCA CCGAGTAGAG AATGTGGCTT CATCTCTCGG CCCCATGCGC 360
 TGGTGGCAGT CCGAGATGA TGTAAACCT GTCTCTCTGC AGCTGGACCT GGACAGGAGA 420
 30 TTCCAGCTT AAGAAGTCAT GATGGAGTTC CAGGGGCCCA TGCCCGCCG CATGCTGATT 480
 GAGCGCTCCT CAGACTTCGG TAAGACCTGG CGAGTGTACC AGTACCTGGC TGCCGACTGC 540
 AACTCCACTT TCCTTCGGGT CCGCCAGGGT CGGCTCAGA GCTGGCAGGA TGTTCGGTGC 600
 CAGTCCCTGC CTCAGAGGCC TAATGCACGC CTAAATGGGG GGAAGGTCCA ACTTAACCTT 660
 35 ATGGATTTAG TGTCTGGGAT TCCAGCAACT CAAAGTCAAA AAATTCAGA GGTGGGGGAG 720
 ATCAAACT TGAGAGTCAA TTTCACACAG CTGGCCCTGT TGCCCAAG GGGCTACAC 780
 CCTCCAGCG CCTACTATGC TGTGTCCAG CTCCTCTGC AGGGGAGCTG CTTCTGTAC 840
 GGCCATGCTG ATGCTCTGCG ACCCAAGCCT GGGGCTCTG CAGGCCCTTC CACCGCTGTG 900
 CAGGTCCACG ATGTCTGTGT CTGCCAGCAC AACACTGCG GCCCAATTG TGAGCGCTGT 960
 40 GCACCTTCT ACTCAACCG CCCTCTGGGA CCGGCGAGG GCCAGGACGC CCATGAATGC 1020
 CAAAGTGGC CACTCAATGG GCACTCAGAG ACATGTCTT TGAACCCGC TGTGTTTGC 1080
 GCCAGCCAG GGCATATGG AGGTGTGTGT GACAAATGCC GGGACCAAC CAGAGGCAAG 1140
 AACTGTGAGC GGTGTCTAGT GCACTATTTC CGGAACCGGC GCCCGGAGC TTCCATTGAG 1200
 GAGACTGCA TCTCTGCGA GTGTGATCG GATGGGGCAG TGCCAGGGC TCCTGTGTAC 1260
 45 CCAGTGGCG GCGAGTGTGT GTGCAAGGAG CATGTGCAAG GAGAGCGCTG TGACCTATGC 1320
 AAGCGGGCT TCACTGGACT GACCTAGGCC AACCCGAGG GCTGCCACG CTGTGACTGC 1380
 AACATCCTG GGTCCCGAG GACATGCCG TGTGACGAG AGAGTGGCG CTGCCTTGT 1440
 CTGCCCAAG TGTGGGTTC CAATGTGAC CAGTGTGCT CCTACACTG GAAGCTGGCC 1500
 AGTGGCCAGC GCTGTGAACC GTGTGCTGC GACCCGACA ACTCCCTCA GCGCACAGT 1560
 50 CAACCACTC ACAGGCGAGT GCCCTGTGCG GAAGGCTTTG GTGGCCTGAT GTGACGCGCT 1620
 GCAGCCATG CCGAGTGTTC AGACCGGACC TATGGAGAG TGGCCACAGG ATGCCGAGCC 1680
 TGTGACTGT ATTTCCGGG AACAGAGGCG CCGGCGTGG ACAGGCGATC AGCGCGCTGC 1740
 CTCGTGCGCC CTGCTGTGAC CGGGCCCGCC TGTGAOCAGT GCCAGCGAG CTACTGCAAT 1800
 CGCTACCCG TGTGCTGCG CTGCCACCT TGTCTCCAGA CCTATGATG GGAOCTCGG 1860
 55 GAGCAGGCCC TGCGCTTGG TAGACTCCG AATGCCACCG CCAGCCTGTG GTGAGGCGCT 1920
 GGGCTGGAG ACCGTGGCT GGCCTCCCG ATCTTAGATG CAAAGAGTAA GATTGAGCAG 1980
 ATCCAGGCA TCTCTAGCAG CCCCAGCTC ACAGAGCAG AGGTGGCTCA GGTGGCCAGT 2040
 GCCATCTCT CCCCAGGCG AACTCTCCAG GGCCTGCAG TGGATCTGCC CCTGGAGGAG 2100
 GAGACGTGT CCCCCTGAG AGACCTGAG AGTCTTGACA GAGGTCTCAA TGTGTCTCT 2160
 60 ACTATGTAT AGAGGAGAG GAGCAGTTT GAAAAATAA GCAGTGTGA TCCTTCAGGA 2220
 GCTTCCGGA TGTCTAGCAC AGCCTAAGAG CAGTCAAGCC AGGCTGCTCA GCAAGTCTCC 2280
 GACAGCTGCG GCCTTTTGA CCAGCTCAG GACAGCGGA GAGAGGCAGA GAGGCTGGT 2340
 CGGCAAGCG GAGGAGGAG AGGCAAGCG AGCCTCAAG TGTGCGCCT GAGGCTGGAG 2400
 ATGTCTTGT TGCCTGACCT GACACCCAC TTCAACAAGC TCTGTGGCAA CTCAGGCAG 2460
 65 ATGGCTTGA CCCCATATC ATGCCCTGCT GAGCTATGTC CCAAGACAA TGGCAGAGC 2520
 TGTGGCTCC GCTGACGGG TGTCTTCCC AGGCGCGGT GGGCCTTCTT GATGGCGGG 2580
 CAGGTGGCTG AGCAGCTGCG GGGCTTCAAT GCCCAGCTCC AGCGACACG GCAGATGAT 2640
 AGGGCAGCG AGGAATCTGC CTCACAGAT CAATCCAGT CCCAGCGCTT GAGACCCAG 2700
 GTGAGCGCA GCGCTCCCA GATGGAGGA GATGTGAGC GCACAGGCT CCTAATCCAG 2760
 70 CAGGTCCGG ACTTCTAAC AGACCCGAC ACTGATGAG CCACTATCCA GAGGTGAGC 2820
 GAGGCGTGC TGGCCCTGT GCTGCCACA GACTCAGCTA CTGTTCTGCA GAGATGAAT 2880
 GAGATCCAG CCAATGCGC CAGGCTCCCG AACGTGGACT TGGTGTCTG CCAGACTCAG 2940
 CAGGACATG CCGGTGCCG CCGGTGCGG GCTGAGGCTG AGGAAGCCAG GAGCGAGCC 3000
 CATGCACTG AGGGCAGGT GGAAGATGT GTTGGGAACC TGCGGCAAGG GACAGTGGCA 3060
 75 CTGAGGAGG CTCAGGACAC CATGCAAGC ACCAGCCCT CCTTCTGCT TATCCAGGAC 3120
 AGGTCTGCT AGGTTCAGCA GGTACTGCG CCAGCAGAAA AGCTGTGTG AAGCATGACC 3180
 AAGCGCTGG GTGACTTCTG GACACGGAT GAGGAGCTCC GCCACCAAG CCGGCGAGCAG 3240
 GGGCAGAGG CAGTCCAGGC CCAGCAGCTT CGCGAAGGT CCAGCGAGCA GGCATTGAGT 3300
 GCGCAAGAG GATTTCAGAG AATAAACA AAGTATGCT AGTTCAGGA CCGGTGGGT 3360
 80 CAGAGTCCA TGCTGGTGA GCGAGGTGCC CCGATCCAGA GTGTGAAGAC AGAGGCGAG 3420
 GAGCTGTTG GCGAGACAT GAGATGATG GACAGGATGA AAGACATGGA GTTGGAGCTG 3480
 CTGCGGGCA GCGAGGCAT CATGCTGCC TCGCGGAGC TGACAGGACT GGAGAAGCGT 3540
 GTGGAGCAGA TCCGTGACCA CATCAATGG CCGTGTCTT ACTATGCCAC CTGCAAGTGA 3600
 TGTACAGCT TCCAGCCGT TGCCCACTC ATCTGCCGCC TTTGCTTTG GTTGGGGGCA 3660
 GATTGGGTTG GAATGCTTTC CATCTCCAG AGACTTTCAT GCAGCTTAA GTACAGCTG 3720

GACCACCCCT GGTGTGTAGC TAGTAAGATT ACCCTGAGCT GCAGCTGAGC CTGAGCCAAAT 3780
 GGACACAGTTA CACTTGACAG ACAAGATGG TGGAGATTGG CATGCCATTG AAACCTAAGAG 3840
 CTCTCAAGTC AAGGAAGCTG GGCTGGGCAG TATCCCCCGC CTTTAGTTCT CCACCTGGGGA 3900
 GGAATCTTGG ACCAAGCACA AAAACTTAAC AAAAGTGATG TAAAAATGAA AAGCCAAATA 3960
 AAAATCTTTG G

Seq ID NO: 320 Protein sequence
 Protein Accession #: NP_000219.1

1 11 21 31 41 51
 MRPFLLCPA LPGLLHAQQA CSRGACYPPV GDLLVGRTRF LRASSTCGLT KPETYCTIYVG 60
 EWQMKCKCD SRQPHNYYSR RVENVASSSG FMRWWSQND VNPVSLQLDL DRRFQLQEVN 120
 MEFGQPMFAG MLIERSSDFG KTWVYQYLA ADCTSTFFRV RQGRPQSNOD VRCQSLPQRP 180
 NARLNGGKVG LNLMDLVSGI PATQSQKIQE VGEITNLRVN FTRLAFVPQR GYHPPSATYA 240
 VSQRLQGSC FCHGHADRC A PKPGASAGPS TAVQVHDVCV CQHNTAGPNC ERCAFFYNNR 300
 FWRPAEQDA HECQRCDNCG HSETHCFDPA VFAASQGAYG GVCNCRDIT EGKNCERCQL 360
 HYFNRNRPGA SIQETCISCE CDFDGAIVPA FCDPVTGQCV CKEHVQGERC DLCKPGFTGL 420
 TYANPQCHER CDNLILGSRR DMPCDEESGR CLCLPNVVG P KCDQCAPYHM KLASGQGCPEP 480
 CACDFHNSPO PTQPVHRAV PCREGFGGLM CSAALIRQCP DRTYGDVATG CRACDCDFRG 540
 TEGGPGCKAS GRCLCRPGLT GPRCDQCQRG YCNRYPCVA CHPCPTVDA DLREQALRFG 600
 LRLNATASLW SGPGLDGRGL ASRILDASKS IEQIRAVLSS PAVTSQSVAG VASAILSLRR 660
 TLQGLQLDLP LSEETLSLPR DLESLEDSFN GLLTMYQRKE EQFEKISSAD PSGAFRMLST 720
 AYEQSAQAAQ QVSDSSRLLD QLRDSRRRAE RLVRQAGGGG GTGSPKLVAL RLEMSLFDL 780
 TPTFNKLGQN SRQMACTPIS CPBELCQQLN GTACGSRRCR VLPRAGGAFI MAGQVAEQLR 840
 GFNAQLQRT R QMLRAAEESA SQIOSSAQLR ETQVSASRSQ MEEDVRRTRL LIQQVDFLIT 900
 DFDIDAATIO EVSEAVLAW LPTDSATVLQ KMNEIQALAA RLNPVLDVLS QTKQDIARAR 960
 RLQAEAREAR SRAHAEVGGV EDVGNLRQG TVALQEAQDT MQGTSRSLRL IQDRVAEVQQ 1020
 VLRPAEKIVT SMTKQLGDFW TRMEELRHQA BQQGAERVA QQLAEGASEQ ALSAQEGFER 1080
 IKQKYAELKD RLQGSMLGE QGARIQSVKT HABELFGETM EMMDRKNOME LELLRGSQAI 1140
 MLRSADLTGL EKRVEQRDHF INGRVLYYAT CK

Seq ID NO: 321 DNA sequence
 Nucleic Acid Accession #: NM_001944.1
 Coding sequence: 84...3083

1 11 21 31 41 51
 TTTTCTTAGA CATTAACTGC AGACGGCTGG CAGGATAGAA GCAGCGGCTC ACTTGGACTT 60
 TTTCACCCAG GAAATCAGAG ACAATGATGG GGCCTCTCCC CAGAACTACA GGGGCTCTGG 120
 CCATCTTGTG GGTGGTCATA TTGGTTCTAG GAGAAATGCG AATAGAGACT AAAGGTCAAT 180
 ATGATGAGA AGAGATGACT ATGCAACAG CTAAAGAAG SCARAAACGT GAATGGGTGA 240
 AATTTCGCAA ACCCTGCAGA GAGGAGAAG ATAACTCAA AAGAAACCCA ATTGCCAAGA 300
 TTACTTCAGA TTACCAAGCA ACCCAGAAA TCACCTACCG AATCTCTGGA GTGGGAATCG 360
 ATCAGCGGCC TTTTGGAAATC TTTGTTGTG ACAAAAACAC TGGAGATATT AACATAACAG 420
 CTATAGTCGA CCGGAGGAA ACTCCAAGCT TCCTGATCAC ATGTCCGGCT CTAATGCC 480
 AAGSACTAGA TGTAGAGAA CCACTTATAC TAACGGTTAA AATTTTGGAT ATTAATGATA 540
 ATCCTCCAGT ATTTTCACAA CAAATTTTCA TGGGTGAAAT TGAAGAAAT AGTGCCCTCA 600
 ACTCACTGTT GATGATACATA AATGCCACAG ATGCAGATGA ACCAAACCCAC TTGAATTTCTA 660
 AAATTGCCCTT CAAATTTGTC TCTCAGGAAC CAGCAGGCAC ACCCATGTC CTCCTAAGCA 720
 GAAACACTGG GGAATCCGT ACTTTGACCA ATCTCTTGA CCGAGAGCAA GCTAGCAGCT 780
 ATCGTCTGGT TGTAGTGGT GCAGACAAAG ATGGAGAGG ACTATCAACT CNAATGTGAAT 840
 GTAATATTAA AGTGAAGAT GTCAACGATA ACTTCCCAAT GTTTAGAGAC TCTCAGTATT 900
 CAGCACGTAT TGAAGAAAT ATTTTAAGTT CTGAATTACT TCGATTTCAA GTAACAGATT 960
 TGGATGAAGA GTACACAGAT AATTGGCCTG CAGTATATTT CTTTACCTCT GGAATGAAG 1020
 GAAATTTGGT TGAATACAA ACTGATCTTA GAACATAAG AGGCATCTCT AAAGTGGTGA 1080
 AGGCTCTAGA TTATGAACAA CTACAAAGCG TGAACTTAG TATTGCTGTC AAAAAACAA 1140
 CTGAATTTCA CCAATCAGTT ATCTCTGAT ACOGAGTTCA GTCAACCCCA GTCAACATTC 1200
 AGGTAATAAA TGTAAAGAAA GGAATTGCAT TCGTCTCTGC TTCCAAGACA TTTACTGTGC 1260
 AAAAAAGCAT AAGTAGCAAA AAATTGGTGG ATTATATCTT GGGAAACATAT CAAGCCATCG 1320
 ATGAGGACAC TAACAAAGCT GCCTCAATG TCAATATGT CATGGGACGT AACGATGGTG 1380
 GATACCTAAT GATTGATTCA AAAACTGCTG AATCAAAAT TGTCAAAAT ATGAACGAG 1440
 ATTCTACTTT CATAGTTAAC AAAACAATCA CAGCTGAGT TCTGGCCATA GATGAATACA 1500
 CGGTAATAAC TTCTACAGGC ACAGTATATG TTAGAGTACC CGATTTCAT GACAAITGTC 1560
 CAACAGCTGT CCTCGAAAAA GATGCAGTTT GCAGTTCTTC ACCTTCCGTG GTTGTCTCG 1620
 CTAGAACACT GAATAATAGA TACACTGGCC CCTATACATT TGCACTGGAA GATCAACCTG 1680
 TAAAGTTGCC TGCCGTATGG AGTATACAA CCTCAATGC TACCTCGGCC CTCTCAGAG 1740
 CCCAGGAACA GATACTCTT GAGGTATACC ACATCTCCCT GGTACTTACA GACAGTCAGA 1800
 ACAATCGGTG TGAGATGCCA CCGAGCTTGA CACTGGAAGT CTGTCAGTGT GACAAACGGG 1860
 GCATCTGTGG AACTTCTTAC CCAACCAAA GCGCTGGGAC CAGGTATGOC AGGCCGCACT 1920
 CAGGGAGGCT GGGGCTCTGC GCCATCGGCC TGCTGTCTCT TGGTCTCTCT CTGCTGTCTG 1980
 TGGCCCCCTT CTGCTGTGTT ACCTGTGACT GTGGGGCAGG TTCTACTGGG GGAATGALAG 2040
 GTGGTTTTAT CCCAGTTCTT GATGGCTCAG AAGGAACAA TCATCAGTGG GGAATTGAAG 2100
 GAGCCCATCC TGAAGACAAG GAAATCACA ATATTGTGT GCTTCTCTGA ACAGCCAAAT 2160
 GAGCGGATT CTGGAAGT TCTGAAGTT GTACAAATAC GTATGCCAGA GGCACAGCGG 2220
 TGGAGGACAC TTCAGGAATG GAAATGAOCA CTAGCTTGG AGCAGCCACT GAATCTGGAG 2280
 GTCTGCGAG CTTTGCACAA GGGACAGTGT CAGGAGCTGC TTCAGGATTC GGAGCAACCA 2340
 CTGGAGTTGG CMTCTGTTC TCAGGGCAGT CTGGAACCAT GAGAACAAAG CATTCCTACT 2400
 GAGGAACCAA TAGGACTAC GCTGATGGGG CGATAAGCAT GAATTTCTCT GACTCTACT 2460
 TTTCTCAGAA AGCATTTGCC TGTGCGGAG AAGACGATGG CCAGGAAGCA AATGACTGCT 2520
 TGTGATCTTA TGAATATGAA GGGCGAGATG CCACTGCTTC TCTGTGGGC TCCGTGGGTT 2580
 GTTGCACTTT TATTCTGAT GACCTGGATG ACAGCTTCTT GGACTCACTT GGACCCAAAT 2640
 TTAATAAACT TGCAGAGATA AGCCTTGGTG TTGATGGTGA AGGCAAGAA GTTCAGCCAC 2700
 CCTCTAAGA CAGCGTTAT GGGATTGAAT CCTGTGGCCA TCCCATAGAA GTCCAGCAGA 2760

CAGGATTGTG TAAGTGCCAG ACTTTGTCAG GAAGTCAAGG AGCTTCTGCT TTGTCCGCT 2820
 CTGGGTCTGT CCAGCCAGCT GTTTCATCC CTGACCCTCT GCAGCATGGT AACTATTAG 2880
 TAACGGAGAC TTACTGGGCT TCTGGTTCCC TCGTGCAACC TTCCACTOCA GGCITTGATC 2940
 CACTTCTCAC ACAAAATGTG ATAGTGACAG AAAGGGTGAT CTGTCCCATP TCCAGTGTTC 3000
 CTGGCAACCT AGCTGGCCCA ACGCAGCTAC GAGGGTCACA TACTATGCTC TGTACAGAGG 3060
 ATCCTTGCTC CGGTCTAATA TGACCAGAAT GAGCTGGAAT ACCACACTGA CCAATCTGG 3120
 ATCTTTGGAC TAAAGTATTC AAAATAGCAT AGCAAAGCTC ACTGTATTGG GCTAATAATT 3180
 TGGCACTTAT TAGCTTCTCT CATAAACTGA TCACGATTAT AATTAATG TTTGGGTTCA 3240
 TACCCCAAAA GCAATATGTT GTCACTCCTA ATTCTCAAGT ACTATTCAAA TTGTAGTAAA 3300
 TCTTAAAGTT TTTCAAAACC CTAATAATCAT ATTCCG

Seq ID NO: 322 Protein sequence
 Protein Accession #: NP_001935.1

1 11 21 31 41 51
 MMGLFPRTTG ALAIFVVVIL VHGLRIETK GQYDEEEMTM QQAKRROKRE NVKFAKPCRE 60
 GEDNSKRNPI AKITSYQAT QKITTYRISGV GIDQPPFGIF VVDKNYGDIN ITAIVDREET 120
 PSPLITCRAL NAQGLDVEKF LILTVKILDI NDNPPVFSQQ IFMGEIERNS ASNSLVMILN 180
 ATDADEPNEL NSKIAFKIVS QEPAGTFMEL LSRNTGEVRT LINSLEDREQA SSYRLVVSQA 240
 DKDGEGLSTQ CECNIVKQDV NDNFPMFRDS QYSARIEENI LSSELLRFQV TDLDEEYTDN 300
 WLAVYFFTSQ NBNWNEFIQT DPRTNBGLK VVKALDYEQL QSVKLSIAVK NKAEPHQSVI 360
 SRYRVQSTFP TIQVINVREG IAFRPASKTF TVQKGISSKK LVDYILGTQ AIDEDTNKAA 420
 SNVRYVMGRN DGGYLMIDSK TABIKFVKNM NRSTFIVNK TITAEVLAI D EYTGKTSIGT 480
 VYVRVDFND NCPTAVLEKD AVCSSPSVV VSARTIANRY TGPYTFALD QPVKLEPAVNS 540
 ITTLNATSAI LRQSQIPPG VYHISLVLD SQNRCSEMPR SLTLEVCQCD NRGICGTSYP 600
 TTSPTGTRYGR PHSGRLGPRA IGLLLGLLLL LLLAPLLLLT CDCGAGSTGG VTGGPIPVFD 660
 GSEGTIHQWG IEGAHPEDEKE ITNICVPPVT ANGADFMESS EVCTINTYARG TAVEGTSQME 720
 MTTKLGAATE SGGAAGFATG TVSGAASGFG AATGVGICSS QGSGTMRTRH STGGTNKDYA 780
 DGAISMNFDL SYFSQKAFAC AEDDDGQEAN DCLLIYDNBS ADATGSEFVS VGCCSFIAAD 840
 LDDSPDLDEL PKPKKLAEIS LGVDGEGKEV QPFSKDSGYG IESCGHPLEV QQTGFVKCQT 900
 LSGSQGASAL SASGSVQPAV SIPDPLQHGK XLVTETYBAS GSLVQPSLAG FDFLLTQNV 960
 VTERVICPIS SVPCNLGAPT QLRGSHMLC TEDPCSRLL

Seq ID NO: 323 DNA sequence
 Nucleic Acid Accession #: XM_058069.2
 Coding sequence: 1..1413

1 11 21 31 41 51
 ATGAAGTTTC TTCTAATACT GCTCCTGCAG GCCACTGCCT CTGGAGCCTT TCCCTGAAAC 60
 AGCTCTACAA GCCTGGAAAA AATAAATGTG CTATTTGGTG AAAGATACTT AGAAAAATTT 120
 TATGGCCTTG AGATAAACAAC ACTTCCAGTG ACAAAAATGA AATATAGTGG AAACCTAATG 180
 AAGGAAAAAA TCCAAAGAAAT GCAGCACTTC TTGGGTCTGA AAGTGACCGG GCACTGAGAC 240
 ACATCTACCC TGGAGATGAT GCACSCACCT CGATGTGGAG TCCCCGATGT CCATCATTTT 300
 AGGGAATGTC CAGGGGGGCC CGTATGGAGG AAACATTATA TCACCTACAG AATCAATAAT 360
 TACACACCTG ACATGAACCG TGAGGATGTT GACTACGCA TCCGGAAGC TTTCCAGTA 420
 TGGAGTAATG TTACCCCTCT GAAATTCAGC AAGATTAACA CAGGCATGGC TGACATTTTG 480
 GTGGTTTTTG CCCGTGGAGC TCATGGAGAC TTCCATGCTT TTGATGSCAA AGGTGGAAAT 540
 CTAGCCCATG CTTTGGGACC TGGATCTGGC ATGGGAGGGG ATGCACATT CTGATGAGAC 600
 GAATYCTGGA CTACACATTG AGGAGGCACA AACTGTGTC TCACTGCTGT TCACGAGATT 660
 GGCCATTCCT TAGGTCTTGG CCACTCTAGT GATCCAAAGC CGGTAATGTT CCCCACCTAC 720
 AATATGTTG ACATCAACAC ATTTGCGCTC TCTGCTGATG ACATACGTTG CATTCAGTCC 780
 CTGTATGGAG ACCCAAAAGA GAACCAACGC TTGCCAAATC CTGACAATTC AGAACAGAT 840
 CTCGTGACCC CCAATTGGAG TTTTGATGCT GTCACTACCG TGGGAAATAA GATCTTTTTC 900
 TTCAAAGACA GGTCTCTCTG GCTGAAGGTT TCTGAGAGAC CAAAGACCAAG TGTTAATTTA 960
 ATTTCTTCTT TATGGCCAAC CTGCGCATCT GGCATTGAAG CTGCTTATGA AATTGAAGCC 1020
 AGAAATCAAG TTTTCTTTT TAAAGATGAC AAATAGTGGT TAATTAGCAA TTTAAGACCA 1080
 GAGCCAAATP ATCCCAAGAG CATACATTCT TTTGGTITTC CTAACTTTGT GAAAAAATP 1140
 GATGCAGCTG TTTTAAACCC ACGTTTITAT AGGACCTACT TCTTTGTAGA TAACAGTAT 1200
 TGGAGGTATG ATGAAAGGAG ACAGATGATG GACCCGTGTT ATCCCAAACT GATTACCAAG 1260
 AACTTCCAAG GAATCGGGCC TAAAATTGAT GCAGTCTTCT ACTCTAAAAA CAAATACTAC 1320
 TATTTCTTCC AAGGATCTAA CCAATTGAA TATGACTTCC TACTCCAAG TATCACCAAA 1380
 ACACGAAAA GCAATAGCTG GTTTGGTGT TGA

Seq ID NO: 324 Protein sequence
 Protein Accession #: P39900

1 11 21 31 41 51
 MKFLILLILQ ATASGALPLN SSTSLKNNV LFGERYLEKF YGLEINKLPV TMKYSGNLM 60
 KEKIQEMQHF LGLKVTGQLD TSTLEMMHAP RGVVDVHHF REMEGGVVNR RHYITYRLNN 120
 YTPDMNREDV DYAIRKAPQV WSNVTPLKFS KINTGMADIL VVFARGAGD FHAFDGGGI 180
 LAHAFPGSGG IGGDAHFDED EFWTTHSGGT NLFLTAVHEI GHSLGLGHS DPKAVMFFTY 240
 KYVDINTFEL SADDIRGLQS LYGPKNOR LNPINSBPA LCDPNLSFDA VTTVGNKLF 300
 FKDRFWMKLV SERPKTVNL ISSLNPILPS GIEAAVEIEA ENQVFLKDD KYNLISNLRP 360
 BPNYKSIHS FGPPNPFVKI DAAVFNRFY RTYFFVDNQY WYDERRQMM DRGYPKLITK 420
 NFQGLGPKID AVFISKNKY YFPQGSNQFE YDFLLQRITK TLKSNWFGC

Seq ID NO: 325 DNA sequence
 Nucleic Acid Accession #: NM_024423.1
 Coding sequence: 64..2590

1 11 21 31 41 51

	GGCAGGCTCTC	GCTCTCGGCA	CCCTCCCGGC	CCCCCGGTTT	TCCTGGCCCT	CCCCGGCATC	60
	CCGATGGCCG	CCGCTGGGCC	CCGCGCTGCC	GTGCGCGGAG	CCGCTGCGCT	GCATCTGCTG	120
	CTGACCCCTC	TGATCTTCAG	TCGTGATGGT	GAAGCCTGCA	AAAAGGTGAT	ACTTAATGTA	180
5	CCTTCTAAAC	TAGAGGCAGA	CAAAATAATT	GGCAGAGTTA	ATTTGGGAAG	GTGCTTCAGG	240
	TCTGCAGACC	TCATCCGGTC	AAGTGATCCT	GATTTCCAGG	TTCTAAATGA	TGGGTCACTG	300
	TACACAGCCA	GGGCTGTGTC	GCTGTCTGAT	AAGAAAAGAT	CATTTACCAT	ATGGCTTTCT	360
	GACAAAAGGA	AACAGACACA	GAAAGAGGTT	ACTGTGCTGC	TAGAACATCA	GAAGAAGGTA	420
	TGGAAGACAA	GACACACTAG	AGAAACGTGT	CTCAGGCGTG	CCAAGAGGAG	ATGGGCCACT	480
10	ATTCCTTGCT	CTATGCAAGA	GAATTCCTTG	GGCCCTTTCC	CATTGTTTCT	TCAACAAGTT	540
	GAATCTGATG	CAGCACAGAA	CTATACCTGC	TTCTACTCAA	TAAGTGGACG	TGGAGTTGAT	600
	AAAGAACCTT	TAAATTTGTT	TTATATAGAA	AGAGACACTG	GAATCTATT	TTGCACTCGG	660
	CCTGTGGATC	TGAAGAATA	TGATGTTTTT	GATTTGATTG	CTTATGCGTC	AACCTGCAGT	720
	GGATATTGAG	CAGATCTGCC	CCTCCCACTA	CCCATCAGGG	TAGAGGATGA	AAATGACAA	780
15	CACCCCTGTT	TCACAGAGGC	AATTTATAAT	TTTGAAGTTT	TGGAAGTAG	TAGACCTGGT	840
	ACTACAGTGG	GGCTGGTTTG	TGCCACAGAC	AGAGATGAAC	CGGACACAA	GCATACGGGC	900
	CTGAATATCA	GCATTTTGCA	CGACAGACCA	AGGTCACTTG	GGCTCTTTTC	TGTGCATCCG	960
	AGCACAGGCG	TAAATCCAC	AGTCTCTCAT	TATTTGGACA	GAGAGGTTGT	AGACAAGTAC	1020
20	TCATTGATAA	TGAAGTACA	AGACATGGAT	GGCCAGTTTT	TTGGATTTGAT	AGGCACATCA	1080
	ACTTGATACA	CTATGCAAGA	GAATTCAAAT	GATAATGCAC	CCACTTTTCA	ACAAAATGCT	1140
	TATGAAGCAT	TTGTAGAGGA	AAATGCATTC	AATGTGGAAA	TCCTACGAAT	ACCTATAGAA	1200
	GATAAGGATT	TAAATTAAC	TGCCAATTTG	AGAGTCAATT	TTACCATTTT	AAAGGGAAT	1260
	GAATAATGAG	TAAATTAAC	CAGCACAGAC	AAAGAACTA	ATGAAGGTGT	TCCTTTCTGT	1320
25	GTAAAGCCAC	TGAATTTATG	AGAAACCGT	CAAGTGACCC	TGGAATTTGG	AGTAACAAAT	1380
	GAAGCCGCTT	TTGCTGAGAG	TATCCGAGA	GTGACAGCCT	TGAACAGAGC	CTTGGTTTAC	1440
	GTTCATGTGA	GGGATCTGGA	TGAGGGGCTT	GAATGCACTC	CTGCAGCCCA	ATATGTGCGG	1500
	ATTAAAGAAA	ACCTTAGCAGT	GGGGTCAAG	ATCAAGGCTT	ATAAGGCATA	TGACCCCGAA	1560
	AAATAGAAAT	GCAATTTGTT	AAGGTACAAA	AAATTTGATG	ATCCTTAAAG	TTGGATTCAC	1620
30	ATTGATGAAA	TTTCCGGGTC	AATCATAACT	TCCAAATCCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCAAAATAT	AGTTGTATTA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
	ACTGGAACAC	TTGCTGTGAA	CATTGAAGAT	GTAAATGATA	ATCCACCAGA	AATACTTCAA	1800
	GAATATGTAG	TCATTTTGCA	ACCAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	TCCATGGAGC	TCCATTTTAT	TTCAATTTGC	CCAATACCTC	TCCAGAAATC	1920
35	AGTAGACTGT	GGAGCCCTAC	CAAGTTAAT	GATACAGCTG	CCGCTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGAAAT	TACCATTCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
	GCAACAAAAT	TATTCAGAGT	TAACTCTGTG	GAATGTACTC	ATCCAACTCA	GTGTCTGTGG	2100
	ACTTCAAGGA	GTACAGGAGT	AATACCTGGA	AAATGGGCAA	TCCTTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCCTTTCTGT	ATTGCTAAT	TTAGTATGTG	GAGTTTITGG	TGCAACTAAA	2220
40	GGGAAACGTT	TTCTGAGAGA	TTTAGCACAG	CAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
	CCTGAGAGAT	TTGCTGATGT	CTTGGCCAT	GGATTTATGA	CCCAACTAC	CAACAACTCT	2340
	AGCCAGGTTT	TTTGTGTATC	TATGGGATCA	GGATGAAAA	ATGGAGGGCA	GGAAACCAAT	2400
	GAATATGATG	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCTCGAGCT	CCCTGAGGGG	AGGACACACG	GAGGTGGACA	ACTGCAGATA	CACCTTACTG	2520
45	GAGTGGCACA	GTTTACTTCA	ACCCGCTCTC	GGTGAAGAA	CCATTAGAGG	ACACACTGGT	2580
	TAAAAATTA	ACATAAGAA	AATTCATCG	ATGTAATCAG	AATGAAGACC	GCATGCAATC	2640
	CCAAGATTAT	GTCTCTCACT	ATACTATGA	GGGAAGGAGA	TCTCCAGCTG	GTCTCTGTGG	2700
	CTGCTGCACT	GAAAGCAGG	AAGAAGATGG	CCTTGACTTT	TAAATAAT	TGGAACCCAA	2760
	ATTTATTACA	TGATCAGAA	CATGCACAAA	GAGATAATGT	CACAGTGCTA	CAATTAGGTC	2820
50	TTTGTGACAG	ATTCCTGAGG	TTTCCAAAA	TAAATATTGA	AAGTTCAATT	TCAACTATGA	2880
	TGTAATGATG	GATTTTITTC	TCAATTTTGA	ATTATGCTAC	TCACCAATTT	ATATTTTAA	2940
	AGCCAGTTGT	TGCTTATCTT	TTCCAAAAAG	TGAAAAATGT	TAAACAGAC	AACCTGTTAA	3000
	TCTCAAACTC	CAGCACTGGA	ATTAAGGCTT	CTAAGGCATC	TGCTCTTTT	TTTTTTTACG	3060
	GATATTTTAG	TAAATAATAT	GCTGGATAAA	TATTAGTCCA	ACAATAGCTA	AGTTATGCTA	3120
55	ATATCACAAT	ATTATGATAT	CACCTTAAGT	GATAGTTTAA	AAAATAAACA	AGAAATATTG	3180
	AGTATCACTA	TGTGAAGAAA	GTTTGGGAAA	AGAAACAAATG	AAGACTGAAT	TAAATTAATA	3240
	ATGTTGACGC	TCATAAGAAA	TTGGGACTCA	CCCTACTGTC	ACTACCAAT	TCATTTGACT	3300
	TTGGAGGCAA	AATGTGTTGA	AGTGCCCTAT	GAAGTAGCAA	TTTTCTATAG	GAATATAGTT	3360
	GGAAATAAAT	GTGTGTGTGT	ATATATTAT	TAACTAATGC	AATATTTAAA	AGAAATGAG	3420
60	AACAAAGAGG	AAATGGGTAA	AAACTTGAAA	TGAGGCTGGG	GTATAGTTTG	TCCTACAAAT	3480
	GAATAAGAGG	AGAGCTTCTT	AGGCTGGGCT	TCCTAAATGC	TGCATTATAA	CTGAGTCTAT	3540
	GAGGAAATAG	TTCTGTGCTA	ATTTGTGTAA	TTTGTTTAAA	ATTGTAAATA	AAATAAACTT	3600
	TTCTGGTTTC	TGTGGGAGGG	AAATAGGGAA	TCCAATGGAA	CAGTAGCTTT	GCTTTGCAAT	3660
	CTGTTTCAAG	ATTTCTGCTAT	CCACAAGTTA	GTAGCAAACT	GGGGAATACT	CGCTGCAAGT	3720
65	GGGGTTCCTT	GCTTTTGTGT	AGCAAGGGTC	CAGAGATGAG	GTGTTTTTTT	CGGGGAGCTA	3780
	ATAACAAAAA	CATTTTAAAA	CTTACCTTTA	CAGAGGTTAA	ATCCTCTATT	GCTGTTTCTA	3840
	TTCTCTCTTA	TAGTGACCAA	CATCTTTTAA	ATTAGATGCC	AAATAACCAAT	GTCTCTCTAG	3900
	AGTTTAGAGG	CTAGAGGGAG	CTGAGGGGAG	GAICTTACTG	AAAGCACCTT	GGGGAGATTG	3960
	ATGTTCTCTA	AACCTAAGCC	CCACAACTT	GACACCTGAT	CAGGTCTGGG	AGCTACAAAA	4020
70	TTTCAATTTT	CCTCTCACTG	CCCTCTCTCT	GAGTGGCAAT	GGCCTGATTC	AAGGAAGGCC	4080
	AGGCTTGTG	GGGCCCCCTC	TTTGGGCTTT	CTGCTAAAGC	AACACCTCCA	GCAGAGATTC	4140
	CTTAAAGTGA	CTCCAGGTTT	TCCACCATCC	TTGAGGTGA	ATTAATTTT	AATCAGTTTG	4200
	CTTTCTCCAG	AGAAATTTTA	AAATAATAGA	AGAAATAGAA	ATTTTGAATG	TATAAAGAAA	4260
	AAAGATCAAG	TGTGCAATTT	AGAACAGAGG	GAACCTTGGG	AGAAAGCAGC	CCAAGTAGGT	4320
75	TATTTGTACA	GTGCAAGGGC	AACAGGAGAA	TGCAAGGCTT	CAAGGCGCAG	GAGAGGCCAC	4380
	AAGGAATATG	GTGGGGAGTA	AAAGCAACAT	CGTCTGCTTC	ATACCTTTTC	CTAGGCTTGG	4440
	CATGCTCTTT	TCCTTTCTCA	GGCCAAATGC	AACCTGCCAT	TGAGTCCGGT	GAGGGATCAG	4500
	CCAACCTCTT	CTCTATGGCT	CACCTTATTT	GGAGTGAGAA	ATCAAGGAGA	CAGAGCTGAC	4560
	TGCATGATGA	GTCTGAGGCG	ATTTCAGGAA	TGAGCCTGAA	CTGGTTGTGC	AGAACAAACA	4620
80	AGGCATTCAT	GGGAATTTGT	GTATTCCTTC	TGCAAGGCTC	CTTCTGGGCA	CTAAGAAGGT	4680
	CTATGAATTA	ATGCGCTATC	TAAATTTCTG	ATTATTTCTT	ACATTTTCTG	TTTTCTAATT	4740
	TGACCCCTAAA	ATCTATGTGT	TTTAGACTTA	GACTTTTAT	TGCCCCCTCC	CCCTTTTTTT	4800
	TTGAGACGGA	GTCTCGCTCT	GAGGCACAGG	CTGGAGTGCA	GTGGCTCCGA	TCTCTGCTCA	4860
	CTGAAAGCTC	CGCTCCCGG	GTTCATGCCA	TTCTCTGCCC	TCAGGCTCCT	GAGTAGCTGG	4920
	GACTACAGGC	GCCCAACACC	AGGCCCGGCT	AATTTTITGT	ATTTTAAATA	GAGACGGGGT	4980

5 TTTCACTGTGT TAGCCAGGAT GGTCTCGATC TCTTGACCTC GTGATCCGCC TGCCTCGGCC 5040
 TCCCAAGATG CTGGGATTAC AGGCATGACC CACCGCTCCC GGCCCTGTGT TCCGTTTAAA 5100
 GTCGTCTTCT TTTAATGTAA TCATTTTGAA CATGTGTGAA AGTTGATCAT ACGAATTGGA 5160
 TCAATCTTGA AATACTCAAC CAAAGACAG TCGAGAGGCC AGGGGGAGAA AGAAGCTCAGG 5220
 GCACAAAATA TTGGTCTGAG AATGGAATTC TCTGTAGCC TAGTTGCTGA AATTTCTGTC 5280
 TGTAACCAAG AGCCAGTTTT ATCTAACCGC TACTGAAACA CCCACTGTGT TTTGCTCACT 5340
 CCCACTCACC GATCAAAACC TGCTACCTCC CCAAGACTTT ACTAGTGCCG ATAACTTTTC 5400
 TCAAAGAGCA ACCAGTATCA CTTCCTGTGT TATAAAACCT CTAACCATCT CTTTGTCTCT 5460
 TGAACATGCT GAAAACCAACC TGCTCTGCAT GTATGCCCGA ATTTGTAAAT CTTTCTCTCT 5520
 10 AATGAAAATA TTAATTTTAG GGATTCATTT CTATATTTTC ACATATGTAG TATTATTATT 5580
 TCCCTATATG TGTAAAGTGA AATTTATGTT ATTTGAGTGT GCAAGAAAAT ATATTTTAA 5640
 AGCTTTTCAAT TTTCCCCCAG TGAATGATTT AGAATTTTTT ATGTAAATAT ACAGAATGTT 5700
 TTTTCTTACT TTTATAAGGA AGCAGCTGTC TAAAATGCAG TGGGGTTTGT TTTGCAATGT 5760
 TTTAAACAGA GTTTTAGTAT TGCTATTAAA AGAAGTTACT TTGCTTTTAA AGAAGCTTGG 5820
 15 CTGCTTAAAA TAACTGAAAA TTGGATGCAT AAGATAATAT TTACAGATGT GGGGAGATGT 5880
 AATAAAACAA TATTACTTGG GCTGCTTAAA ATAAGCAAAA ATTTGATGCA TAAAGTAATA 5940
 TTTACAGATG TGGGAGATG TAATAAAACA ATATTAACCT GGTTTCTTGT TTTTGTCTGA 6000
 TTTAGAGATT AATAAATCTT AAGATGATCA CTTTGCAAAA TTATGCTTAT GGCCTGCCATG 6060
 GAAATAGAAA TACTCAATTA TGTCTTTGTT GTATTAATGG GGAATATTTT GGCAATGTT 6120
 20 TCAATATCAA TTAATTCGACA TCATTAAATAT ATATTGTAAT GTTGGGAAGA GATCACTATT 6180
 TTGAAGCACA GCCTTACAGA TCAGTATCTA TGATACATAT GTATAATAAA TTTTGTATCGG 6240
 GTATTAAAGG TATTAGAAGG TGGTTATAAT TGCAGAGTAT TCCATGAATA GTCACTGAC 6300
 ACAGGGGTTT TACTTTGAGG ACCAGTGTAG TCAAGGGAAA ACATGAGTTA AAAAGAAAAG 6360
 CAGGCAATAT TGCAGCTCTG ATCTGCCCAC TTACAGGATA GATAATGCCCT GAACCTTAAT 6420
 25 GACAAGATGA TCCAACCATTA AAGGTGCTCT GTGCTTCACA GTGAATCTTT TCCCATGCA 6480
 GGAGTGTGCT CCCCTACAAA CGTTAAGACT GATCATTTC AATATCTATT AGCTATATCA 6540
 AAAGCCTTAC ATTTTAAAT AGGTTGAACC AAAATTTCAA TTCCAGTAAC TTCTATTGTA 6600
 ACCATTATTT TGTGTATGT CTTCAAGAAAT GTTCATTGGA TTTTGTGTTG TAAATAGTAA 6660
 ATACCGGATA CATTTCCAGT GTCCCTCAGT ATTGATTGG TTGAATATTG GGTCTAATG 6720
 30 GTTGAAGAGC ATGGACACTA GAGCCAGAAAT GCTTGGATAT GAATCCCTGA TCTGTCACTT 6780
 ACTTCTGTGT GACCTTTGAA AGGCTACTTA TTTCTCTCT TAGCTTTCTC ATTAATAATCA 6840
 ATGAACAATG CCAGCCTCAT GGGTTGTGTT AATGATTAAA TTAGTTAATA TACCTAAAGT 6900
 ACATAGAACCA CTGCTGTCAC ATAGTAAAAG AATTATAGT GTGAGGTAGT TGGTAAATTT 6960
 35 ATGTAGTTGG ATATACTACC GAACAATATC TAATCTCTTT TTAGGGAAAT AAAGTTTGTG 7020
 CATATATATA ATCCCAAAC ATG

Seq ID NO: 326 Protein sequence
 Protein Accession #: NP_077741.1

40 1 11 21 31 41 51
 MAAAGPRRSV RGAVCLHLLT TLVIFSRDGE ACKKVIINVP SKLEADKIIG RVNLEECFRS 60
 ADLIRSDSDP FRVLNDGSVY TARAVALSDK KRSFTIWLSD KKKQTQKEVT VLLHQKKVS 120
 45 KTRHRTYVL RRAKRRNAPI PCSMQENSLG PFPLFLQVVE SDAAGNYTVF YSLSGRGVDK 180
 EPLNLFYIER DTGNLFCTRP VDREYDVFD LIAVASTADG YSADLPLPLP IRVEDENDNH 240
 FVTEALYNF EYLESRPRT TVGVVCAADR DEPDTHMTRL KYSILQQTFR SPGLFSVHPB 300
 TGVITTVSRY LDREVDVKYS LIMKVQDMG QFFGLIGTST CLITVTDEND NAPTFRQAY 360
 EAFVSENAFN VEILRIPIED KOLINTANWE VNFTILKME NGHFKISTDK EINEGVLSVV 420
 50 KPLNYEENRQ VMLSIGVWNE APPARDIPRV TALNRALVTV HVRDLDEGPR CTPAAQVRI 480
 KENLAVGSKI NGYKAYDPEN KNGGKLYKK LEDPKGWTI DEISGSIITS KILDREVEPT 540
 KENLYNITVL AIDKDRDST GTLAVNLEDV NUNPPEILQE YVVICPKMG YTDILAVDPD 600
 EPVHGAFPFY SLPTNSPERS RLWSLTKVND TAARLSYQKN AGFQRYTIP I TVKDRAGQAA 660
 TKLLEVNLCB CTHFTQCRAT SRSTGVILOK WAILAILLGI ALLFSLVILX VCVFPGATKG 720
 55 KRFPEDLAQQ NLIIISNTBAP GDDRVCBANG FMTQTNNSS QGFCGTMGSG MKNGGGQETLE 780
 MMKGGAQETLE SCRGAQHHEH LDCRQGETE VDNCRITYSE WHSFTQPRLG EESTRGHTG

Seq ID NO: 327 DNA sequence
 Nucleic Acid Accession #: NM_001941.2
 Coding sequence: 64..2754

60 1 11 21 31 41 51
 GGCAGGTCCT GCTCTCGGCA CCGTCCCGGC GCCCGCGTTC TCGTGGCCCT GCCCGCATC 60
 CCGATGGCCG CCGCTGGGCC CCGCGCTGCC GTGCGCGGAG CCGTCTGCCT GCACTCTCTG 120
 65 CTGACCCCTCG TGATCTTCAG TCGTGATGGT GAAGCCTGCA AAAAGGTGAT ACTTAATGTA 180
 CCTTCTAAAC TAGAGGCAGA CAAATAAAT GGCAGAGTTA ATTTGGAAGA GTGCTTCAGG 240
 TCTGCAGACC TCATCCGGTC AAGTGAOCT GATTTCAGAG TTCTAAATGA TGGGTCACTG 300
 TACACAGCCA GCGCTGTTGC GCTGTCTGAT AAGAAAAGAT CATTTACCAT ATGGCTTTCT 360
 70 GACAAAAGGA AACAGACACA GAAAGAGGTT ACTGTCTGTC TAGAACATCA GAAGAAGGTA 420
 TCGAAGACAA GACACACTAG AGAAACTGTT CTCAGGCGTG CCAAGAGGAG ATGGGCACCT 480
 ATTCTCTGCT CTATGCAAGA GAATPCCTTG GGCCTTTCC CATGTGTTCT TCAACAAGTT 540
 GAATCTGATG CAGCACAGAA CTATACTGTC TTCTACTCAA TAAGTGGAGC TGGAGTTGAT 600
 AAAGAACCTT TAAATTTGTT TTATATAGAA AGAGACACTG GAAATCTAAT TTGCACTCGG 660
 75 CCTGTGATC GTGAAGAATA TGATGTTTTT GATTGTATTG CTATGCGTC AACTGCAGAT 720
 GGATATTGAG CAGATCTGCC CCTCCACTA CCATCAGGG TAGAGGATGA AATGACAAAC 780
 CACCTGTGTT TCACAGAAGC AATTATTAAT TTTGAAGTTT TGGAAAGTAG TAGACCTGGT 840
 ACTACAGTGG GGGTGGTTTG TGCCACAGAC AGAGATGAAC CGGACACAAT GCATACGGCG 900
 CTGAATACA GCATTTTGC A GACAGACCA AGGTACCTG GGCTCTTTTC TGTGCTCCC 960
 AGCAGAGCGG TAATCACCAC AGTCTCTCAT TATTGGACA GAGAGGTTGT AGACAAGTAC 1020
 80 TCAATGATAA TGAAGATACA AGACATGGAT GGCCAGTTT TTGATTTGAT AGGCACATCA 1080
 ACTTGTATCA TTAGCATAAC AGATTCAAT GATAATGCAC CCACCTTCAG ACAAATGCT 1140
 TATGAAGCAT TTGTAGAGAA AATGCAATC AATGTGGAAT TCTTACAAAT ACCTATAGAA 1200
 GATAAGGATT TAATTACAC TGCCATTTGG AGAGTCAAT TTACCATTTT AAAGGGAAT 1260
 GAAATGGAC ATTTCAAAT CAGCACAGAC AAAGAACTA ATGAAGGTGT TCTTCTGTT 1320

	GTAAAGCCAC	TGAATTATGA	AGAAAAACCGT	CAAGTEAACCC	TGGAAATTTGG	AGTAAACAAT	1380
	GAAGCGCCAT	TGTCTAGAGA	TATTCGCCAG	GTGACAGCCT	TGAACAGAGC	CTTGGTTACA	1440
	GTTCATGTGA	GGGATCTGGA	TGAGGGGCCCT	GAATGCACCT	CTGCGAGCCCA	ATATGTGCGG	1500
5	ATTAAAGAAA	ACTTAGCAGT	GGGGTCAAAG	ATCAACGGCT	ATAAGGCATA	TGACCCCGAA	1560
	AATAGAAATG	GCAATGGTTT	AAGGTACAAA	AAATTGCATG	ATCCTAAAGG	TTGGATCACC	1620
	ATTGATGAAA	TTTCAGGGTC	AATCATAACT	TCCAAATCC	TGGATAGGGA	GGTTGAAACT	1680
	CCCAAAATG	AGTTGTATAA	TATTACAGTC	CTGGCAATAG	ACAAAGATGA	TAGATCATGT	1740
	ACTGGAACAC	TGCTGTGAA	CATGGAAGAT	GTAAATGATA	ATCCACCAGA	AACTACTCAA	1800
10	GAATATGTAT	TCAATTGCAA	ACCAAAATG	GGGTATACCG	ACATTTTAGC	TGTTGATCCT	1860
	GATGAACCTG	CCATGGAGC	TCCATTTTAT	TTGAGTTTGC	CCAATACTTC	TCCAGAAATC	1920
	AGTAGACTGT	GGAGCCTCNC	CAAAGTTAAT	GATACAGCTG	CCCGTCTTTC	ATATCAGAAA	1980
	AATGCTGGAT	TTCAAGAATA	TACCATTCCCT	ATTACTGTAA	AAGACAGGGC	CGGCCAAGCT	2040
	GCAACAAAAT	TATTGAGAGT	TAATCTGTGT	GAATGTACTC	ATCCAACTCA	GTGTGTTGCG	2100
15	ACTTCAAGGA	GTACAGGAGT	AATACTTGGA	AAATGGGCAA	TCCCTGCAAT	ATTACTGGGT	2160
	ATAGCACTGC	TCCTTTCTGT	ATTGCTAACT	TTAGTATGTG	GAGTTTTTGG	TGCAACTAAA	2220
	GGGAAACGTT	TTCTGGAAGA	TTTAGSCACAG	CAAACTTAA	TTATATCAAA	CACAGAAGCA	2280
	CCTGGAGAGC	ATAGAGTGTG	CTCTGCCAAT	GGATTTATGA	CCCAACTTAC	CAACAACCTT	2340
	AGCCAGGTCT	TTTGTGTGAC	TATGGGATCA	GGAAATGAAA	ATGGAGGGCA	GGAAACCAAT	2400
20	GAATATGATG	AAGGAGGAAA	CCAGACCTTG	GAATCCTGCC	GGGGGGCTGG	GCATCATCAT	2460
	ACCTCGAGCT	CTGCGAGGG	AGGACACACG	GAGGTGGACA	CTGCAGATA	CACCTACTCG	2520
	GAGTGGCACA	GTTTTACTCA	ACCCCGTCTC	GGTGA AAAAT	TGCATCAGTG	TAATCAGAAAT	2580
	GRAGACCGCA	TGCCATCCCA	AGATTATGTC	CTCACTTATA	ACTATGAGGG	AAGAGGATCT	2640
	CCAGCTGGTT	CTGTGGGCTG	CTGCAGTGAA	AAGCAGGAG	AAGATGGGCT	TGACTTTTTA	2700
25	AATAATTTGG	AACCCAAAT	TATTACATTA	GCAGAAGCAT	GCACAAGAG	ATAATGTCTC	2760
	AGTGTACAG	TTAGTCTTGT	GTGAGACATT	CTGGAGGTTT	CCAAAATATA	TATTGTAAAG	2820
	TTCAATTTCA	ACATGTATGT	ATATGATGAT	TTTTTCTTCA	ATTTTGAAT	ATGCTACTCA	2880
	CCAAATTTATA	TTTTTAAAGC	CAGTTGTGTC	TTATCTTTTC	CAAAAAGTGA	AAAATGTTAA	2940
	AACAGACAAC	TGGTAATATC	CAAACTCCAG	CACCTGGAAT	AAGGTCTCTA	AAGCATCTGC	3000
30	TCTTTTCTTT	TTTACGGAT	ATTTTAGTAA	TAAATATGCT	GGATAAATAT	TAGTCCAACA	3060
	ATAGCTAAGT	TATGCTAATA	TCACATTATT	ATGTATTAC	TTTAAGTGAT	AGTTTAAAAA	3120
	ATAAACAGAA	AATATTGAGT	ATCCTATGTT	GAAGAAAGTT	TTGGAAGAGA	AACATGAAG	3180
	ACTGAATTA	ATTAAAAATG	TTGCAGCTCA	TAAAGAAATG	GGACTCACCC	CTACTGCACCT	3240
	ACCAAAATCA	TTTGAATTTG	GAGGCAAAAT	GTGTTGAAGT	GCCCTATGAA	GTAGCAATTT	3300
35	TCTATAGGAA	TATAGTTGGA	AATAAATGTG	TGTGTGTATA	TTATTATTAA	TCAATGCAAT	3360
	ATTTAAATG	AAATGAGAA	AAAGAGGAAA	ATGGTAAAAA	CTTGAATGA	GGCTGGGGTA	3420
	TAGTTTGTCC	TACAATAGAA	AAAAGAGAGA	GCTTCTTAGG	CCTGGGCTCT	TAAATGTCTC	3480
	ATTATACTG	AGTCTATGAG	GAATAGTTCT	CTGTCCAATT	TGTGTAAATT	GTTTAAAAAT	3540
40	GTAAATAAAT	TAACTTTTTC	TGTTTCTGT	GGGAAGGAAA	TAGGGGAATCC	AATGGAACAG	3600
	TAGCTTTGCT	TTGCAGTCTG	TTTCAAGATT	TCTGCATCCA	CAAGTTAGTA	GCAAACTGGG	3660
	GAATCTGCG	TGCAGCTGGG	GTTCCTGCT	TTTTGGTAGC	AAGGGTCCAG	AGATGAGGTG	3720
	TTTTTTTGG	GGAGCTAATA	ACAAAACAT	TTTAAAACTT	ACCTTTACTG	AAGTTAAATC	3780
	CTCTATTGCT	GTTCCTATTC	TCTCTTATAG	TGACCAACAT	CTTTTAAAT	TAGATCCAAA	3840
45	TAACTATGTC	CTCTCATGAGT	TTAGAGGCTA	GAGGGAGCTG	AGGGGAGGAT	CTTACTGAAA	3900
	GCACCTGGG	GAGATTGATT	GTCCCTTAAAC	CTAAGCCCCA	CAAACTGTAC	ACCTGATCAG	3960
	GTCTGGAGC	TAATGAGTAT	CATTTTCTC	CTCACTGCC	TTCTTCTGAG	TGGCATTTGC	4020
	CTGAATCAAG	GAAAGCCAGG	CCTTGTGGGC	CCCTTCTTT	CGGCTTCTG	CTAAGCAAC	4080
	ACCTCCAGCA	GAGATTCCCT	TAACTGACTC	CAGGTTTTCC	ACCATCCTTC	AGCGTGAAT	4140
50	AATTTTAAAT	CAGTTTCTCT	TCTCCAGAGA	AAITTTAAAA	TATAGAGAGA	AATAGAAAT	4200
	TTGAATGAT	AAAGAGAAAA	GATCAAGTTG	TCAITTTAGA	ACAGAGGAAA	CTTTGGGAGA	4260
	AAGCAGCCCA	AGTGGTTAT	TGTACAGTC	AGAGGGCAAC	AGGAAGATGC	AGGCCTTCAA	4320
	GGGCAAGGAG	AGGCCACAAG	GAATATGGGT	GGGAGTAAAA	GCAACATCGT	CTGCTTCAAT	4380
	CTTTTCTCTA	GGCTTGGCAC	TGCCCTTTCC	TTTCTCAGGC	CAATGGCCAC	TGCCATTTGA	4440
55	GTCCGGTGG	GGATCAGCCA	ACCTCTTCTC	TATGGCTCAC	CTTATTTGGA	GTGAGAAAT	4500
	AAGGAGACAG	AGCTGACTGC	ATGATGATTC	TGAAGGCATT	TGCAGGATGA	GCCTGAACTG	4560
	GTGTGTGACA	ACAAACAAAG	CATTCTATGG	AATGTTGTGA	TTCTTCTGTC	AGCCCTCTCT	4620
	CTGGGCTCTA	AGAAGGTCTA	TGAATTAAAT	GCCTATCTAA	AATTTCTGAT	TATCTCTACA	4680
	TTTTCTGTTT	TCTAATTTGA	CCCTAAATC	TATGTGTTTT	AGACTTAGAC	TTTTTATTGC	4740
60	CCCCCCCCC	TTTTTTTTTG	AGACGGAGTC	TGCTCTGAC	GCACAGGCTG	GAGTGCAGTG	4800
	GCCTCCGATC	CTGCTCACTG	AAAGCTCCGC	CTCCCGGTT	CATGCCATTC	TCCCTGCTCA	4860
	GCCTCTGAG	TAGCTGGGAC	TACAGGCGCC	CACCACCAAG	CCCGGCTAAT	TTTTTGTATT	4920
	TTTAAATAG	ACGGGGTTTC	ACTGTGTTAG	CCAGGATGGT	CTCGATCTCC	TGACCTCGTG	4980
	ATCCGCTGCT	CTCGGCTTCC	CAAAGTGCTG	GGATTAACAG	CATGACCCAC	CBCTCCCGGC	5040
	CTTGTCTTCC	GTTTAAAGTC	GTCTCTTTT	AATGTAATCA	TTTTGAACAT	GTGTGAAAGT	5100
65	TGATCATACG	AATGGATCA	ATCTTGAAT	ACTCAACCAA	AAGACAGTCC	AGAAGCCAGG	5160
	GGGAGAAAGA	ACTCAGGGCA	CAAAATATTG	GTCTGAGAAAT	GGAAATCTCT	GTAAAGCTAG	5220
	TTGCTGAAT	TTCTGCTGTT	AACAGAAAGC	CAGTTTTATC	TAAAGGCTAC	TGAACACACC	5280
	ACTGTGTTTT	GCTCACTCCC	TCACTCACCG	ATCAAAACCT	GCTACCTCCC	CAAGACTTTA	5340
	CTAGTGCCGA	TAACTTTCTT	CARAGAGCAA	CCAGTATCAC	TTCCCTGTTT	ATAAAACCTC	5400
70	TAACTATCTC	TTTGTCTTTT	GAACATGCTG	AAAACCACTT	GGTCTGCTAG	TATGCCCCBA	5460
	TTTGTAAATC	TTTTCTCTCA	AATGAAATTT	TAAATTTAGG	GATTCAITTC	TATATTTTCA	5520
	CATATGTAGT	ATATATATTT	CCTTATATGT	GTAAAGTGAA	ATTTATGGTA	TTTGAGTGTG	5580
	CAAGAAATA	TATTTTTTAA	GCTTTCAAT	TTCCCCAGT	GAATGATTTA	GAATTTTTTA	5640
	TGTAAATATA	CAGAAATGTT	TTTCTTACT	TTATAGGAA	GCAGCTGTCT	AAAATGCAGT	5700
75	GGGGTTTTGT	TGCAATGTT	TTAAACAGAG	TTTATGATTT	GCTATTAATA	GAATGTACTT	5760
	TGCTTTTAAA	GAACCTTGGC	TGCTTAAAT	AAGCAAAAT	TGGATGCATA	AAGTAATATT	5820
	TACAGATGTG	GGGAGATGTA	ATAAAACAA	ATTAACCTGG	TTTCTGTGTT	TGCTGTATT	5880
	TAGAGATTAA	ATAATCTTAA	GATGATCACT	TTGCAAAAT	ATGCTTATGG	CTGGCATGGA	5940
	AATAGAAATA	CTCAATATG	TCTTGTGTT	ATTAATGGGG	AATATTTTGG	ACATATTTTC	6000
	ATTATCAAT	TGTCAGATC	ATTAATATAT	ATGTATATGT	TGGGAAGAGA	TCACTATTTT	6060
80	GAAGCACAGC	TTTCTGATG	AGTATCTATG	ATACATATGT	ATAATAAAT	TTGATCGGGT	6120
	ATTAAGAGTA	TAGAAGGTTG	GTTATAATTT	CAGATATTC	CATGAATAGT	ACACTGACAC	6180
	AGGGTTTTTA	CTTTGAGGAC	CAGTGTAGTC	AAGGGAAAC	ATGAGTTTAA	AAGAAAGCA	6240
	GGCAATATTG	CAGTCTTGAT	TCTGCCACTT	ACAGGATAGA	TAAATGCTGA	ACTTTAATGA	6300
	CAAGATGATC	CAACATAAA	GGTGTCTGTT	GCTTCACAGT	GAATCTTTTC	CCCATGCAGG	6360

5 AGTGTGCTCC CCTACAAACG TTAAGACTGA TCATTTCAAA AATCTATTAG CTATATCAAA 6420
 AGCCITACAT TTATATATAG GTTGAACCAA AATTTCATTT CCAGTAACTT CTATTGTAAC 6480
 CATATATTTT GTGATGTGCT TCAAGAAATGT TCATTGGATT TTGTTTGTGA ATAGTAAAAAT 6540
 ACCGGATACA TTTCACGGGT CCTTCAGTAT TGATTGGGT GAATATTGGG TCATAATGGT 6600
 TGAGAAGCAT GGACACTAGA GCCAGAAATGC TTGGATATGA ATCCTGGATC TGTCACTTAC 6660
 TTCGTGTGA CCTTGAAG GCTACTTATT TCCTCTCTTA GCTTCTCAT TAAATCAAT 6720
 GAACAATGCC AGCCTCATGG GGTGTGTGAA TGATTAAAT AGTTAATATA CCTAAGTAC 6780
 ATAGAACACT GCTGCACAT AGTAAAAGAA TTATAAGTGT GAGGTAGTTG GTAAAATTAT 6840
 GTAGTTGGAT ATACTACCGA ACAATATCTA ATCTCTTTT AGGGAATAA AGTTTGTGCA 6900
 10 TATATATAAT CCCGAAACAT G

Seq ID NO: 328 Protein sequence
 Protein Accession #: NP_001932.1

15 1 11 21 31 41 51
 | | | | |
 MAAAGFRFRS RGAVCGLLLL TLVIFSRDGE ACKKVILNVP SKLEADKIIG RVNLEECFRS 60
 ADLIRSDPD FRVLNDGSEVY TARAVALSDK KRSEFTINLSD KRKQTQKEVT VLLEHQKKVS 120
 20 KIRHTRETVL RRAKRWAWPI PCSMQENSLG FFLFLQQVE SDAAQNYTVF YSISORGVDK 180
 EPLNLFYIER DTGNLFCTRP VDREYDVED LIAYASTADG YSADLEPLFP IRVEDENDNE 240
 PVFTEAIYNF EVLESSRPGI TVGVVCATDR DEPDTHMTRL KYSILQQTFR SPGLFSVHPS 300
 TGVITTVSHY LDREVVDKYS LIMKVQDMDG QFFGLIGTST CIITVDSND NAPTFRQWY 360
 EAFVSENAFN VEILRIPIED KDLINTANWR VNFILKENE NGHPKISIDK ETNEGVLVV 420
 25 KPLNYEENRQ VNLIEGVNNE AFFARDIPRV TALNRALVTV HVRDLDEGPE CTPAAQYVRI 480
 KENLAVGSKI NGKAYDPEN RNNGRLRYKK IHDPKGNITI DEISGSIITS KILDRREVTF 540
 KNELYNITVL AIDKDRDSCT GLAVNIEDV NDNPEEILQE YVVICPKMG YTDILAVDPD 600
 EPVHGAFYF SLFNTSFEIS RLWSLTQVND TAARLSYQKN AGPQEVYTIPI TVKDRAGQAA 660
 TKLRLVNLCE CTHPTQCRAT BRSTGVILGK WAILAILLGI ALLFSVLLTL VCGVFGATKG 720
 KRFPEDLAQQ NLLISNTEAP GDDRVCSANG FMTQTNNSS QGFCGTMSGG MKNGGQETIE 780
 30 MNKGNQTLK SCRGAGHHHT LDCRGGHTE VDMCRYTYSR WBSFTQPRLG EKLHRCNQNE 840
 DRMPQDYVL TYNVEGRGSP AGSVGCCSEK QREGLDPLN NLEPKFITLA EACTKR

Seq ID NO: 329 DNA sequence
 Nucleic Acid Accession #: NM_016583.2
 Coding sequence: 72..842

40 1 11 21 31 41 51
 | | | | |
 GGAGTGGGGG AGAGAGAGGA GACCAGGACA GCTGCTGAGA CCTCTAAGAA GTCCAGATAC 60
 TAAGAGCAAA GATGTTTCAA ACTGGGGGCC TCATTGTCTT CTACGGGCTG TTAGCCCGAGA 120
 CCATGGCCCA GTTGGAGGCG CTGCCCGTGC CCTTGGACCA GACCCGTGCC TTGAATGTGA 180
 ATCCAGCCCT GCCCTTGAGT CCCACAGGTC TTGCAGGAAG CTTGACAAAT GCCCTCAGCA 240
 45 ATGGCTGCTG GTCGGGGGCG CTGTTGGCCA TTCTGGAAAA CCTTCCGCTC CTGGACATCC 300
 TGAAGCCTGG AGGAGGTACT TCTGGTGGCC TCCTTGGGGG ACTGCTTGA AAGSTGACGT 360
 CAGTGATTCC TGGCCTGAAC AACATCATTG ACATAAAGGT CACTGACCCC CAGCTGCTGG 420
 AACCTGGCCT TGTGAGAGC CCTGATGGCC ACCGTCTCTA TGTACCATC CCTCTCGGCA 480
 TAAAGTCCA AGTGAATACG CCGCTGGTCG GTGCAAGTCT GTTGAGGCTG GCTGTGAAGC 540
 TGGACATCAC TGCACAAATC TTAGCTGTGA GAGATAAGCA GGAGAGGATC CACCTGTGCC 600
 50 TTGGTGAAG CACCCATGCC CCGTGAAGCC TGCATAATTC TCTGCTTGAT GGACTTGGCC 660
 CCGTCCCCAT TCAAGTCTT CTGACAGGCC TCACAGGGAT CTGGAATAAA GTCCTGCCTG 720
 AGTTGGTTCA GGGCAACGCG TGCCCTCTGG TCRATGAGGT TCTCAGAGGC TTGACATCA 780
 CCGTGGTGCA TGACATTGTT AACATGCTGA TCCACGGACT ACAGTTTGTG ATCAGAGTCT 840
 AAGCCTTCCA GGAAGGGGCT GGCCTCTGCT GAGCTGCTTC CCAGTGCTCA CAGATGGCTG 900
 55 GCCCATGTGC TGGAGATGGA CACAGTTGCC TTCTCTCGGA GGAADCTGCC CCGTCTCTT 960
 TCCACACAGG CGTGTGTAAC ATCCCATGTG CCTCACTTAA TAAATGGCT CTTCTTCTGC 1020
 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 330 Protein sequence
 Protein Accession #: NP_057667.1

60 1 11 21 31 41 51
 | | | | |
 MFQTGGLIVF YGLLAQTMAQ FGLFVPLDQ TLPLNVNPAI PLSPTGLAGS LTNALSNGLL 60
 SGGLLGLLEN LPLLDILKPG GGTSGGLLGG LLGKVTSVIP GLNNIIDIKV TDPQLLELGL 120
 65 VQSPDGERLY VTPIPLGKLQ VNTPLVGASL LRLAVKLDIT ABLLAVRDKQ ERIHLVLGDC 180
 THSPGSLQIS LLDGLGLPLI QGLLDLSLGI LNKVLPVLVQ GNVCPLVNEV LRGLDITLVH 240
 DIVNMLIHGL QFVIKV

Seq ID NO: 331 DNA sequence
 Nucleic Acid Accession #: NM_004363.1
 Coding sequence: 115..2223

75 1 11 21 31 41 51
 | | | | |
 CTCAGGGCAG AGGGAGGAAG GACAGCAGAC CAGACAGTCA CAGCAGCCTT GACAAAACGT 60
 TCTCGAATC CAAGCTCTTC TCCACAGAGG AGGACAGAGC AGACAGCAGA GACCATGGAG 120
 TCTCCTCGG CCCCTCCCCA CAGATGGTGC ATCCCTCTGGC AGAGGCTCCT GCTCAGAGCC 180
 80 TCACITCAA CCTTCTGGAA CCGGCCACCC ACTGCCAAGC TCACTATTGA ATCCAGGCCG 240
 TCAATGTGC CAGAGGGGAA GAGGTGCTT CTACTGTGCC ACAATCTGCC CCAGCATCTT 300
 TTTGGCTACA GCTGGTACA AGGTGAAAGA GTGGATGGCA ACCGTCAAAT TATAGGATAT 360
 GTRATAGGAA CTCACCAAGC TACCCAGGG CCGCATAACA GTGGTCAAGA GATAATATAC 420
 CCCAATGCAT CCTGCTGAT CCAGAACATC ATCCAGAATG ACACAGGATT CTACACCTTA 480

CACGTCAATA AGTCAGATCT TGTAATGAA GAAGCAACTG GCCAGTTCGG GGTATACCCG 540
 GAGCTGCCCC AGCCCTCCAT CTCAGCAAC AACTCCAAAC CGTGGAGGA CAAGGATGCT 600
 GTGGGCTTCA CCTGTGAACC TGAGACTCAG GACGCAACCT ACCTGTGGTG GGTAAACAAT 660
 CAGGACCTCC CGGTCAATTC CAGGCTGCAG CTGTCCAAATG GCAACAGGAC CCTCACTCIA 720
 5 TTCAATGTCA CAAGAATGSA CACAGCAAGC TACAAATGTG AAACCCAGAA CCCAGTGAGT 780
 GCCAGGCGCA GTGATTCAGT CATCTGTAAT GTCTCTATG GCCCGGATGC CCCACCAATT 840
 TCCCTCTCAA ACACATCTTA CAGATCAGGG GAAATCTGA ACCTCTCTG CCACGCGGCC 900
 TCTAACCCAC CTGCACAGTA CTCTTGGTTT GTCAATGGGA CTTCACAGCA ATCCACCCAA 960
 10 GAGCTCTTTA TCCCAACAT CACTGTGAAT AATAGTGGAT CCTATACGTG CCAAGCCCAT 1020
 AACTCAGACA CTGGCTCTCA TAGGACCACA GTCAAGCAAG TACAGTCTA TGACAGAGCA 1080
 CCAAAACCTT TCATCACCAG CAACAACCTC AACCCCGTGG AGGATGAGGA TGCTGTAGCC 1140
 TTAACCTGTG AACCTGAGAT TCAGAACACA ACCTACCTGT GGTGGGTAAA TAATCAGAGC 1200
 CTCCCGGTCA GTCCCAAGCT GCAGCTGTCC AATGACAACA GGACCCCTAC TCTACTCAGT 1260
 15 GTCAACAGGA ATGATGTAGG ACCCTATGAG TGTGGATCC AGAACGAAT AAGTGTGAC 1320
 CACAGCGACT CTGGCTCTCA TAGGACCACA GTCAAGCAAG TACAGTCTA TGACAGAGCA 1380
 TCTATACCTT ATTACCTGTC AGGGGTGAAC CTCAGCCTCT CCTGCCATGC AGCCTCTAAC 1440
 CCACCTGACG AGTATCTTGT GCTGATTGAT GGAACATCC AGCAACACAC ACAAGAGCTC 1500
 TTTATCTCCA ACATCACTGA GAAGAACAGC GGAATCTATA CCTGCCAGGC CAATAACTCA 1560
 20 GCCAGTGGCC ACAGCAGGAG TACAGTCAAG ACAATCAGAG TCTCTGGGGA GCTGCCCAAG 1620
 CCTCTCATCT CTGGCTCTCA TAGGACCACA GTCAAGCAAG TACAGTCTA TGACAGAGCA 1680
 TGTGAACCTG AGGCTCAGAA CACAACCTAC CTGTGGTGGG TAAATGGTCA GAGCCTCCCA 1740
 GTCAGTCCCA GGCTGAGCTT TCCCAATGGC AACAGGACCC TCACTCTATT CAATGTCA 1800
 AGAAATGACG CAGAGGCTTA TGTATGTGGA ATCCAGAACT CAGTGAAGTC AAGCCGAGT 1860
 25 GACCCAGTCA CCTCTGATGT CCTCTATGAG CCGGACACCC CCATCATTTT CCCCCAGAC 1920
 TCGTCTTACC TTTGGGAGG GAACTCTAAC CTCTCTGCTC ACTCGGCTCT TAACCCATCC 1980
 CCGCAGTATT CTGGGCTTAT CAATGGGATA CCGCAGCAAC ACACACAAGT TCTCTTATC 2040
 CGCAAAATCA CGCCAAATAA TAAAGGGACC TATGCTCTTT TGTCTCTAA CTGGCTACT 2100
 GGCCGCAATA ATTCCATAGT CAAGAGCATC ACAGTCTCTG CATCTGGAAC TCTCTCTGTT 2160
 30 CTCTCAGCTG GGGCCACTGT CGGCATCATG ATTGGAGTGC TGGTGGGGT TGCTCTGATA 2220
 TAGCAGCCCT GGTGTAGTTT CTTCATTTCA GGAAGACTGA CAGTGTGTTT GCTTCTCTCT 2280
 TAAAGCATTT GCACAGCTA CAGTCTAAA TGTCTCTTT ACCAAGGATA TTTACAGAAA 2340
 AGACTCTGAC CAGAGATCGA GACCATCTTA GCCAACATCG TGAAACCCCA TCTCTACTAA 2400
 AAATACAAAA ATGAGCTGGG CTGGTGGGG CACACCTGTA GTCCCAAGTA CTGGGAGGC 2460
 35 TGAGGCGAGA GAATCGCTTG AACCCGGGAG GTGGAGATTG CAGTGAGCTC AGATCGCACC 2520
 ACTGCACTCC AGTCTGGCAA CAGAGCAAGA CTCATCTCA AAAAGAAAAG AAAAGAGAGC 2580
 TCTGACCTGT ACTCTGAAT ACAAGTTCTT GATACCATCG CACTGTCTGA GAATTTCCAA 2640
 AACTTAAATG AACTGAACGA CAGCTTCATG AAACGTCCA CCAAGATCAA GCAGAGAAA 2700
 40 TAATTAATTT CATGGGACTA AATGAACATA TGAGGATTGC TGATCTTTA AATGTCTTGT 2760
 TTCCAGATT TCAGGAACAT TTTTCTCTT TAAGCTATCC ACTCTTACAG CAATTTGATA 2820
 AAATATACTT TTGTGAACAA AAATGAGAC ATTACATT TCTCCCTATG TGGTCTCTCC 2880
 AGACTTGGGA AACTATTAT GAATATTAT ATTGTATGT AATATAGTTA TTGCACAAGT 2940
 TCAATAAAAA TCTGCTCTTT GTATAACGA AAAA

Seq ID NO: 332 Protein sequence

Protein Accession #: NP_004354.1

1 11 21 31 41 51
 | | | | |
 50 MESPSPAPPHR WCIPWQRLLL TASLLTFWNP PTTAKLTIES TPFNVABGKE VLLLVHNLPLQ 60
 HLFQYSWYRG ERVDGNRQII GYVIGTQQAT PGPAYSGRBI IYPNABLLIQ NIIQNDTGFI 120
 TLHVIKSDLV NEETAGQFRV YFELPKPSIS SNNSKPEVDK DAVAFICEPE TQDATYLNWV 180
 NNQSLFVSER LQLSMGNREL TLFNVTRNDT ASYKCBTQNP VSARRSDSVI LNVLYGPDAP 240
 TISPLNTSYR SGHNLNLCH AASNPQAQYS WFWNGTFQQS TQELFIPNIT VNSGSGYTCQ 300
 55 AHNSDTGLNR TIVTITVYA BPPKFFITSN NSNPVEDEDA VALTCERPEIQ NITYLNWVWN 360
 QSLFVSERLQ LNDNRNLTL LSVTRNDVGP YECGIONELS VHSDFVILN VLYGPDPTI 420
 SPSTYYRFG VNLBSLCHAA SNPPQAQYSWL IDGNIQEQET ELFIKITEK NSGLYTCQAN 480
 NSASGHSRTT VKTITVSABL PKPSISNNNS KPVEDKDAVA FTCEPEAQNT TYLNWVNGQS 540
 LPVSPRLQLS NGNRILTLFW VTRNDARAYV CGIQNSVSN RSDFVTLDVL YCPDTPPISP 600
 60 PDSBYLSGAN LNLCHSASN PSPQYSWRIN GIPOQHTQVL FLAKITPNMN GTYACFVSNL 660
 ATGRNNSIVK SITVBSGTS PLSAGATVG IMIGVLVGVA LI

Seq ID NO: 333 DNA sequence

Nucleic Acid Accession #: NM_006952.1

Coding sequence: 11..793

1 11 21 31 41 51
 | | | | |
 65 AATCCCGACA ATGGCGAAAG ACAACTCAAC TGTGTGTGTC TTCCAGGGCC TGCTGATTTT 60
 TGGAAATGTG ATTATGTGTT GTTGGCGCAT TGCCCTGACT GCGGAGTGCA TCTTCTTTGT 120
 70 ATCTGACCAA CACAGCTCTT ACCCACTGCT TGAAGCCACC GACAAAGATG ACATCTATGG 180
 GGCTGCGCTG ATCGGCATAT TGTGGGCTAT CTGCTCTCTC TGCTGTCTG TTTAGGCGAT 240
 TGTAGGCTAT ATGAAGTCCA GCAGGAAAT TCTTCTGGCG TATTTCATTC TGATGTTTAT 300
 AGTATATGCC TTTGAAGTGG CATCTGTGAT CACAGCAGCA ACACAACGAG ACTTTTTCAC 360
 75 ACCCAACTCT TTCTGAAGC AGATGCTAGA GAGGTACCAA AACACAGCC CTCCAAACAA 420
 TGATGACCAG TTGAAAACAA ATGGAGTCCAC CAAAACCTGG GACAGGCTCA TGCTCCAGGA 480
 CAATGCTGCT GGGGTAAATG GTCCATCAGA CTGGCAAAAA TACACATCTG CCTTCCGGAC 540
 TGAGAAATAT GATGCTGACT ATCCCTGGCC TCGTCAATGC TGTGTTATGA ACAATCTTAA 600
 AGAACCTCTC AACCTGGAGG CTGTAAACT AGGCGTGCTT GGTTTTATC ACAATCAGGG 660
 80 CTGCTATGAA CTGATCTCTG GTCCAAATGA CCGACAGGCC TGGGGGGTGG CTTGGTGTGG 720
 ATTTGCCATT CTCTGCTGGA CTTTTGGGT TCTCTGGGT ACCATGTCTT ACTGGAGCAG 780
 AATTGAATAT TAAGAA

Seq ID NO: 334 Protein sequence

Protein Accession #: NP_008883.1

```

1      11      21      31      41      51
|      |      |      |      |      |
5  MAKDNSTVRC FQGLLIFGNV IIGCCGIALT ASCIFFVSDQ HSLYPLLEAT DNDIYGAAN 60
  IGIFVGICLF CLSVLGIVGI MKSSRKILLA YFILMFIVYA FEVASCITAA TQRDPFTPNL 120
  FLKQMLERYQ NNSFENNDDQ WKNNGVTKIW DRLMLQDNCC GVNGPBDWOK YTSAFRTENN 180
  DADYFWRQC  CVMNMLKEPL NLEACKLGVP GPYHNGQCYE LISGPMNRHA WGVAVGFPAI 240
  LCWTFWVLIG TMFYWSRIEY

```

Seq ID NO: 335 DNA sequence
 Nucleic Acid Accession #: NM_002638.1
 Coding sequence: 120..473

```

1      11      21      31      41      51
|      |      |      |      |      |
15 CAATACAGCT AAGGAATAT CCCTTGTAAT TACCCAGAC CGCCCTGGA GCCAGGCCAA 60
  GCTGGACTGC ATAAAGATTG GTATGGCCTT AGCTCTTAGC CAAACACCTT CCTGACACCA 120
  TGAGGCCAGC CAGCTTCTTG ATCGTGTGTG TGTTCCTCAT CGCTGGGACG CTGGTTCTAG 180
  AGGCAGCTGT CACGGGAGTT CCTGTTAAAG GTCAAGACAC TGTCAAAGGC CGTGTTCCTAT 240
  TCAATGGACA AGATCCCBTT AAGGACAAG TTTCAGTTAA AGGTCAAGAT AAAGTCAAAG 300
  CGCAAGAGCC AGTCAAAGGT CCAGTCTCCA CTAAGCCTGG CTCTGCCCC ATTATCTTGA 360
  TCGGTGGCG CATGTTGAAT CCCCTAACC GCTGCTTGA AGATACTGAC TGCCCAAGAA 420
  TCAGAAGTG CTGTGAAGGC TCTTGCGGGA TGGCCTGTTT CFTTCCCGAG TGAAGGAGC 480
  CGGTCCCTGC TGCACCTGTG CCGTCCCGAG AGCTACAGGC CCAATCTGAT CTAAGTCCC 540
  TGCTGCCCTT CCCCTTCCA CACTGTCCAT TCTTCTCC ATTCAGGATG CCCACGGCTG 600
  GAGCTGCCCT TCTCATCCAC TTTCAATAA A

```

Seq ID NO: 336 Protein sequence
 Protein Accession #: NP_002629.1

```

1      11      21      31      41      51
|      |      |      |      |      |
30 MRASSFLIVV VFLLAGTLVL EAAVIGVPEK GQDTVGRVP FNGQDPVKIQ VSVKQDKVK 60
  AQEFVVKFVS TFGSCPIIL IRCAMLNFFN RCLKDTDCPS IKKCCRGSCG MACFVVPQ

```

Seq ID NO: 337 DNA sequence
 Nucleic Acid Accession #: NM_001793.2
 Coding sequence: 71..2560

```

1      11      21      31      41      51
|      |      |      |      |      |
40 AAAGGGGCAA GAGCTGAGCG GAACACCGGC CGCCCGTGGC GGCAGCTGCT TCACCCCTCT 60
  CTCTGCAGCC ATGGGGCTCC CTCTGGGACC TCTGCGGCT CTCTCTCTTC TCCAGGTTTG 120
  CTGGCTCGAG TGGCGCGCCT CGAGCGCGTG CGGGCGGCTC TTCAAGGAGG CTGAAGTGAC 180
  CTGGAGGCG GAGGCGCGCG AGCAGGAGCC CGGCCAGGCG CTGGGGAAGG TATTCTATGG 240
  CTGCCCTGGG CAAAGGCCAG CTCTGTTTAG CACTGATAAT GATGACTTCA CTGTGCGGAA 300
  TGGCGAGACA GTCCAGGAAA GAAGGTCACT GAAGGAAAGG AATCCATTGA AGATCTTCCC 360
  ATCCAAACCT ATCTTACGAA GACACAAGAG AGATTGGGTG GTTGCTCCA TATCTGTCCC 420
  TGAAATGGC AAGGGTCCCT TCCCCAGAG ACTGAATCAG CTCAAGTCTA ATAAAGATAG 480
  AGACACCAAG AATTTCTACA GCATCACGGG GCGCGGGGCA GACAGCCCC CTGAGGGTGT 540
  CTTCGCTGTA GAAAGGAGA CAGGCTGGTT GTTGTGAAT AAGCCACTGG ACCGGGAGGA 600
  GATTGCCAAG TATGAGCTCT TTGGCCACGC TGTGTCAAG AATGGTCCCT CAGTGGAGGA 660
  CCCCATGAC ATCTCCATCA TCGTAGCCA CCAGAATGAC CACAAGCCCA AGTTTACCCA 720
  GGACACCTTC CGAGGGAGTG TCTTAGAGGG AGTCTTACCA GTTACTTCTG TGATGCRGGT 780
  GACAGCCACG GATGAGGATG ATGCCATCTA CAOCTACAAT GGGGTGGTTG CTTACTCCAT 840
  CCATAGCCAA GAACCAAGG ACCACACA GA CCTCATGTTC ACCATTCACC GGAGCACAGG 900
  CACCAACAC GTCTCTCCA GTGGCCTGGA CGGGAAAAA GTCCCTGAGT ACACACTGAC 960
  CATCCAGGCC ACAGACATGG ATGGGACCGG CTCACCACCC ACGGCAGTGG CAGTAGTGGA 1020
  GATCCTGAT GOCATGACA ATGCTCCCAT GTTTGACCCC CAGAAATACG AGGCCCATGT 1080
  GCCTGAGAA GTAGTGGGCC ATGAGGTGCA GAGGCTGACG GTCACTGATC TGGACGCCCC 1140
  CAACACCA GGTGGGGTGG CCACCTACCT TATCATGGGC GGTGAAGAGG GGGACCAATT 1200
  TACCATCACC ACCACCCCTG AGAGCAACCA GGGCATCTG ACAACCAAGG AGGTTTGA 1260
  TTTTGAGGCC AAAAACCAGC ACACCCCTGA CGTTGAAGTG ACCAAGAGG CCCCCTTTGT 1320
  GCTGAAGCTC CCAACCTCCA CAGCCACCAT AGTGGTCCAC GTGGAGGATG TGAATGAGGC 1380
  ACCTGTGTT GTCCACCTCT CCAAGTCTGT TGAGGTCCAG GAGGGCATCC CCACTGGGGA 1440
  GCTGTGTGT GTCTACACTG CAGAAGACCC TGACAGGAG AATCAAGA TCACTACCG 1500
  CATCTTGAGA GACCCAGCAG GTTGGCTAGC CATGACTTCA GACAGTGGGC AGGTACAGC 1560
  TGTGGGCACC CTGACCGGTG AGGATGAGCA GTTGTGAGG AACACATCT ATGAAGTCAT 1620
  GGTCTTGGCC ATGGACAATG GAAGCCCTCC CACCACTGGC ACGGGAACCC TTCTGCTAAC 1680
  ACTGATTGAT GTCAATGACC ATGGCCAGT CCCTGAGGCC CGTCAGATCA CCACTCTGAA 1740
  CCAAGCCCT GTGCGCCAGG TGCAGAACAT CACGGACAAG GACCTGTCTC CCCACACCTC 1800
  CCTTTTCCAG GCCCAGCTCA CAGATGACTC AGACATCTAC TGGACGGCAG AGGTCAACGA 1860
  GGAAGGTGAC ACAGTGGTCT TGTCCCTGAA GAAGTTCCCT AAGCAGGATA CATATGACGT 1920
  GCACCTTTCT CTGTCTGACC ATGGCAACAA AGAGCAGCTG ACGGTGATCA GGGCCACTGT 1980
  GTGCACTGAG CATGACCATG TCGAAACCTG CCCTGGACCC TGGAAAGGGAG GTTTCATCTT 2040
  CCTGTGTCTG GAGGCTGTCT TGGCTCTGCT GTTCTCTCTG CTGTTGCTGC TTTTGTGTGT 2100
  GAGAAGGAG CGGAAGATCA AGGAGCCCTT CCTACTCCCA GAAGATGACA CCGGTGACAA 2160
  CGTCTTCTAC TATGGGGAAG AGGGGGGTGG CGAAGAGGAC CAGGACTATG ACATCACCCA 2220
  GCTCCACCGA GGTCTGGAGG CCAGGCCGGA GGTGTTTCTC CGCAATGAGG TGGCACCAC 2280
  CATCATCCCG ACACCATGAT ACGTCTCTCG GCCAGGCCAC CAGATGAAA TOGGCAACTT 2340
  TATAATTTAG AACCTGAGG CGGCTAACAC AGACCCACCA GCGCCGCCCT ACGACACCTT 2400
  CTTGCTGTTC GACTATGAG GCAGCGGCTC CGACGCGGCB TCCCTBAGCT CCTCACCTC 2460
  CTCGCGCTCC GACCAAGACC AAGATTGAGA TTAICTGAAC GAGTGGGCA GCCCTTCAA 2520
  GAAGCTGGCA GACATGTACG GTGGCGGGA GGACGACTAG GCGGCTGCTC TGCAGGCTG 2580

```

5
10
GGGACCAAAAC GTCAGGCCAC AGAGCATCTC CAAGGGGTCT CAGTTCCCCC TTCAGCTGAG 2640
GACTTCGGAG CTGTGTCAGGA AGTGGCCGTA GCAACTTGGC GGAGACAGGC TATGAGTCTG 2700
ACGTTAGAGT GGTTCCTTCC TTAGCCTTTC AGGATGGAGG AATGTGGGCA GTTTGACTTC 2760
AGCACTGAAA ACCTCTCCAC CTGGGCCAGG GTTGCTCTAG AGGCCAAATT TCAGAGAGCC 2820
TCTTACCTGC CGTAAATGTC TCAACCTGT GTCTGGGCC TGGCCCTGCT GTGACTGACC 2880
TACAGTGGAC TTTCTCTCTG GAATGGAACC TTCTTAGGCC TCCTGCTGCA ACTTAATTTT 2940
TTTTTTAAT GCTATCTTCA AAACGTTAGA GAAAGTCTT CAAAAGTGCA GCCCAGAGCT 3000
GCTGGGCCCA CTGGCCGTCC TGCAATTTCTG GTTTCAGAC CCCAATGCCT CCCATTCCGA 3060
TGGATCTCTG CTTTATTATA CTGAGTGTGC CTAGGTTGCC CCTTATTTT TATTTTCCCT 3120
10 GTTGGTTGC TATAGATGAA GGTGAGGAC AATCGTGTAT ATGTACTAGA ACTTTTAT 3180
TAAAGAACT TTTCCAGAA AAAAA

Seq ID NO: 338 Protein sequence
Protein Accession #: NP_001784.2

15
20
25
30
1
11 21 31 41 51
MGLPRGFLAS LLLLVQCNLO CAASEPCRAV FREAEVTLEA GGAEQEPQQA LGKVFMGCPG 60
QEPALFSTDN DDFTVRNGET VQERRSLKER NPLKIFPSKR ILRRHKRDWV VAPISVPENG 120
KGFFPQRLNQ LKSNKDRDTK IFYSITGPGA DSPPGVFAV EKETGWLLIN KPLDREBIK 180
YELFGHAVSE NGASVEDPMN ISIIVTDQND HKPKFTQDTE RGSVLEGVLE GTSMVQTAT 240
DEDDAIYTYN GVVAYSIESQ BPKDPHDLMP TIHRSTGTIS VISSGLDREK VPEYTLTIOA 300
TDMDDGDSIT TAVAVVEILD ANDNAPMFDP QKYSAHVPEM AVGHEVQRLT VTDLDAPNSP 360
AMRATYLLMG GDDGDHFTIT THPESNQGIL TTRKGLDFEA KQHTLYVEV TNEAPFVLKL 420
PTSTATIVVH VEDVNEAPVF VPPEKVVVEQ EGIPTGEPVC VYTAEDPDKE NQKISYRILR 480
DPAGWLAMD PDSQVITAVGT LDREDEQFVR NNIEYVMVLA DMNGSFFFTG TGTLLLTLD 540
VNDRGPVPEP RQITICNQSP VRQVINITDK DLSPTSPFPQ AQLTDDSDIY WTAEVNBECD 600
TVVLSLKKFL KQDTYIVHLG LSDHGNKEQL TVIRATVDCD HGHVETCPGP WKGGFILPVL 660
GAUALLFL LLVLLLVRLK RKIKFPLLLP EDDTRDNVFI YGEGGGGED QDYDITQLER 720
30 GLEARPEVVL RNDVAPTII TPMYRPRPAN FDEIGNFIE NLKAAITDPT APPYDTLLVF 780
DYEGSGSDAA SLSSLTSSAS DQDQDYDYLN BWGSRPKLA DMVGGCEDD

Seq ID NO: 339 DNA sequence
Nucleic Acid Accession #: Eos sequence
Coding sequence: 1..672

35
40
45
50
1
11 21 31 41 51
ATGAGGCTCC AAAGACCCCG ACAGGCCCGG GCGGCTGGGA GGCGCGCGCC CCGGGGCGGG 60
CGGGGCTCCC CTTACCGGCC AGACCGGGGG AGAGGCGCGC GGAGGCTGCG AAGGTTCCAG 120
AAGGGGCGGG AGGGGGCGCC GCGCGCTGAC CCTCCCTGGG CACCGCTGGG GACGATGGCG 180
CTCTCGGCTT TGCTGCTGGT CGTGGCCCTA CCGCGGGTGT GGACGAGCGC CAACCTGACT 240
GCGAGACAA GAGATCCAGA GGACTCCAG CGAAGGACG AGGGTGACAA TAGAGTGTGG 300
TGTCATGTTT GTGAGAGAGA AAACACTTTC GAGTGCCAGA ACCCAAGGAG GTGCAAAATCG 360
45 ACAGAGCCAT ACTGCTTAT AGCGGCGGTG AAAATATTTC CACGTTTTTT CATGGTTGCG 420
AAGCAGTGT CCGCTGTTTG TGCAGCGATG GAGAGACCCA AGCCAGAGGA GAAGCGGTTT 480
CTCCGGAAG AGCCCATGCC CTCTTTTAC CTCAGTGTGT GTAAAATTTC CTACTGCAAT 540
TTAGAGGGGC CACCTATCAA CTCATCAGTG TTCAAAGGAT ATGCTGGGAG CATGGGTGAG 600
50 AGCTGTGGTG GGCTGTGECT GGCCATCCCT CAGCTGCTGG CCTCCATTGC AGCCGCGCTC 660
AGCCTGTCTT GA

Seq ID NO: 340 Protein sequence
Protein Accession #: Eos sequence

55
60
1
11 21 31 41 51
MRLQRPRQAP AGGRHAPRGG RGSFYRDPG RGARRLRRFQ KGEGGAPRAD PFWAPLGTHA 60
LLALLLVVAL FRVWTDANLT ARQRDPDSQ RTDEGDNKRV CHVCERENTF BCQNPFRCKW 120
TEPYCVIAAV KTFPRFFMVA KQCSAGCAAM BRKPEEKRF LLRBPMPFFY LKCKIRYCN 180
LEGPPINSSV FREYAGSMGE SCGLNLAIL LLLASIAAGL SLS

Seq ID NO: 341 DNA sequence
Nucleic Acid Accession #: XM_035292.2
Coding sequence: 53..1576

65
70
75
80
1
11 21 31 41 51
GCTGCTGGG CCGCGGCTCC CCGGTGTGCC AGGCCCGGCC GTTGCGCAGA GCATGGCGGG 60
TGCGGGCCCG AAGCGCGCGC CGCTAGCGGC GCGCGCGGCC GAGGAGAGAG AAGAGCGCGC 120
GGAGAAGATG CTGGCGGCCA AGAGCGCGGA CGGCTCGGCG CCGGCAGGCG AGGGCGAGGG 180
CGTGACCCCTG CAGCGGACCA TCACGCTGCT CAACGGCGTG GCCATCATCG TGGGGAACAT 240
TATCGGCTCG GGCATCTTGG TGAGCGCCAC GGGCGTGCTC AAGGAGGACG GCTCGCGGG 300
GCTGGCGCTG GTGGTGCGGG CCGGTGTGCC CGTCTTCTCC ATCGTGGGCG CGCTCTGCTA 360
CGCGGAGCTC GGCACACCA TCTCCAAATC GGGCGCGCAC TACGCTTACA TGCTGGAGGT 420
CTACGCTCAT CTGCGCGCTT TCTCAAGCT CTGGATGAG CTGCTCATCA TCCGCGCTTC 480
ATCGCAGTAC ATCGTGGGCC TGGTCTTCGC CACTACCTG CTCAGGCGC TCTTCCCCAC 540
CTGCCCCGTG CCGGAGGAGG CAGCCAGCT CTGCGCGTGC CTCGCGTGC TGCTGCTCAC 600
GGCGGTGAAC TGCTACAGCG TGAAGGCGGC CACCCGCGTC CAGGATGCTT TTGCGCGCGC 660
CAAGCTCCCTG GCGCTGGGCC TGATCATCCT GCTGGGCTTC GTCCAGATCG GAAAGGGTGA 720
75 TGTGTCCAT CTGATGCCCA ACTTCTATT TGAAGGCACC AAACCTGGAT TGGGGAACAT 780
TGTGCTGCA TTATACAGCG GCTCTTTGCG CTATGAGGGA TGGAAATTA TGAATTTGCT 840
CACAGAGGAA ATGATCAACC CCTACAGAAA CCTGCCCTCG GCCATCATCA TCTCCCTGCC 900
CATCGTGAAG CTGGTGTACG TGCTGACCAA CCTGGCTTAC TTCACCAACC TGCTCACCGA 960
80 GCAGATGCTG TGTTCGAGG CCGTGCCGT GGACTTCGGG AACTATCACC TGGGCTCAT 1020

GTCTCGATC ATCCCGTCT TCGTGGGCT GTCTGCTTC GGCTCCETCA ATGGGTCCCT 1080
 GTTCACATCC TCCAGGCTCT TCTTGTGGG GTCCCGGAA GGCCACCTGC CCTCCATCCT 1140
 CTCATGATC CACCCACAGC TCCTCACCCC CGTGCCGTCC CTCGTGTTC CGTGTGTGAT 1200
 GACGCTGCTC TAGCCTTCT CCAAGGACAT CTCTCCGTC ATCAACTTCT TCAGCTTCTT 1260
 CAACTGCTC TGGTGGGCC TGGCCATCAT CGGCATGATC TGGCTGCGCC ACAGAAAGCC 1320
 TGAGCTGAG CGGCCATCA AGGTGAACCT GGCCTGCTC GTGTCTTCA TCCTGGCCTG 1380
 CCTCTTCTG ATCGCGTCT CCTTCAGGA GACACCCGTG GAGTGTGGCA TCGGCTTAC 1440
 CATCATCTTC AGCGGGCTGC CCGTCTACTT CTTCGGGTC TGGTGGAAA ACAAGCCCA 1500
 GTGGCTCTC CAGGGCATCT TCTCCAGAC CGTCTGTGT CAGAAGCTCA TGCAGGTGGT 1560
 CCCCAGGAG ACATAGCCAG GAGGCGAGT GGTGCGGGA GGAGCATGC

Seq ID NO: 342 Protein sequence
 Protein Accession #: XP_035292.2

1 11 21 31 41 51
 | | | | |
 MAGAGPKRRA LAAPAAEKEE EAREKMLAAK SADGSAPAGE GEGVTLQRNI TLINGVAIIV 60
 GTIIGSUIFV TPTGVLEKEG SPGLALVWMA ACGVFSIVGA LCTAELGTIT SKSGGDYAYM 120
 LEVYGSILPAP LKIMIELLII RPSSQYIVAL VFATYLLKPL FPTCPVPEEA AKLVACLCLV 180
 LLTAVNCSYV KAATRVQDAF AAALLLALAL IILGPFVQIG KGDVSNIDPN FSPFGTKLDV 240
 GNIVLALYSY LPAYGWNLYL NFVTEEMINP YRNLPALIII SLPIVTLVYV LTNLAYFTTL 300
 STEQMLSSAA VAVDFGNLYL GVMNIIIVF VGLSCFGSVN GSLFTSSRLF FVGSREGLP 360
 SILSMIHPQL LTPVPSLVPT CVMILLIAYS KDIFSVINEF SFENWLCVAL AIIGMIWLRH 420
 RKPELERPIK VNLALVFFFI LACLFIAVS FWKTPVECGI GPTIILSGLP VYFFGVWWMK 480
 KPKNLQGIIF STTVLCQKLM QVVPQST

Seq ID NO: 343 DNA sequence
 Nucleic Acid Accession #: NM_005268.1
 Coding sequence: 168..989

1 11 21 31 41 51
 | | | | |
 TAAAAAGCAA AAGAATTCGC GCGCGCTCG ACACGGGCTT CCGGAAAAC CTTCCTCGCT 60
 TCTGGATATG AATTCACAGC TGCTTGTCTG GTCTATTGTC CGGCTGCTGG GAGCCAGGAG 120
 AGCCCTGAGG AGTAGTCACT CAGTAGCAGC TGACGCGTGG GTCCACCATG AACTGGAGTA 180
 TCTTTGAGGG ACTCTGAGT GGGGTCAACA AGTACTCCAC AGCTTTTGGG CGCATCTGGC 240
 TGTCTCTGGT CTTCATCTTC CGCGTCTGG TGTACTTGGT GACGCGCAGC CGTGTGTGGA 300
 GTGATGACCA CAAGGACTTC GACTGCAATA CTCGCCAGCC CGGCTGCTCC AACGTCTGCT 360
 TGTATGAGTT CTTCCTCTGT TCCCATGTGC GCCTCTGGGC CCTGCACCTT ATCTTGGTGA 420
 CATGCCCTTC ACTGCTCTGT GTCATGCACG TGGCCTACCG GGAGGTTCAG GAGAAGAGGC 480
 ACCGAGAAGC CCATGGGGAG AACAGTGGGC GCCTCTACCT GAACCCCGGC AAGAAGCGGG 540
 GGGGCTCTG GTGCACATAT GTCTGCAGCC TAGTGTTCAG GCGAGCGGTG GACATCGCCT 600
 TTCTCTATGT GTCCACATCA TTCTACCCCA AATATATCCT CCTCTCTGTG GTCAAGTGCC 660
 ACGCAGATCC ATGTCCCAAT ATAGTGGACT GCTTCATCTC CAGGCCCTCA GAGAAGAACA 720
 TTTTCACCTT CTTCATGGTG GCCACAGCTG CCATCTGCAT CTGTCTCAAC CTCGTGGAGC 780
 TCATCTACCT GGTGAGCAAG AGATGCCACG AGTGCTTGGC AGCAGGAAA GCTCAAGCCA 840
 TGTGCACAGG TCATCAACCC CACGGTAACA CCTCTCTCTG CAACAAGAC GACCTCCTTT 900
 CGGCTGACCT CATCTTTCTG GCTCAGACA GTCATCTCC TCTCTTACCA GACCGCCCC 960
 GAGACCATGT GAAGAAAACC ATCTTGTGAG GGGCTGCTCG GACTGGTCTG GCAGGTGGG 1020
 CTTGATGGG GAGGCTCTAG CATCTCTCAT AGGTGCAACC TGAGAGTGGG GGAGCTAAGC 1080
 CATGAGGTAG GGGCAGGCAA GAGAGAGGAT TCAGACGCTC TGGGAGCCAG TTCTTAGTCC 1140
 TCACTCCAG CCACTGCCCC CAGCTCGACG GCACCTGGCC AGTTCCCCCT CTGCTCTGCA 1200
 GCTCGGTTTC CTTTCTCTAGA ATGGAATAG TGAGGCCAA TGC

Seq ID NO: 344 Protein sequence
 Protein Accession #: NP_005259.1

1 11 21 31 41 51
 | | | | |
 MNNSIFEGLL SGVNYSTAF GRINLSLVFI FRVLVYLVTI ERVNSDDHKO FDCNTRQPGC 60
 SNVCFDEFFP VSHVRLMALQ LILVTCPSSL VVMVAYREV QEKRHREAHG ENSGRILYNP 120
 GKRGGLWMT YVCSLVFKAS VDIAPLYVFH SFYFKYLPP VVKCHADPCP NIVDCFISKP 180
 SEKNIFPLFM VATAACILLL NLVELIYLYS KRCECLAAK KAQAMCTGHH PHGTTSSCKQ 240
 DDLLSGDLIF LGSDSHPPIL PDRPRDHVK TIL

Seq ID NO: 345 DNA sequence
 Nucleic Acid Accession #: NM_002391.1
 Coding sequence: 26..457

1 11 21 31 41 51
 | | | | |
 CGGGCGAAGC AGCGCGGGCA GCGAGATGCA GCACCGAGGC TTCTCTCTCC TCACCTCTCT 60
 CGCCTCTCTG GCCTCACTC CGCGGTCTGC CAAAAGAAA GATAAGGTGA AGAAGGGCGG 120
 CCGGGGGAGC GAGTGCCTTG AGTGGGCTTG GGGGCTTGC ACCCCAGCA GCAAGGATTG 180
 CGCGGTGGGT TTCCGGAGG GCACCTGCGG GGGCCAGACC CAGCGCTTCC GGTGCAAGGT 240
 GCGCTGCAAC TGAAGAAGG AGTTTGAGC CAGCTGCAAG TACAAGTTTG AGAATGGGG 300
 TGGGTGTGAT GGGGGCACAG GCACCAAGT CCGCCAAGGC ACCCTGAAGA AGGCGGCTA 360
 CAGTCTCAG TGCAGGAGA CCATCCGCT CACCAAGCCC TGCAACCCCA AGACCAAGC 420
 AAGGCCAAA GCCAAGAAAG GGAAGGAAA GCACTAGACG CCAAGCCTGG ATGCCAAGG 480
 GCGCTGGTG TCATGAGGG CCTGGCCAGC CCCTCTCTCT CCCAGGCCG AGATGTGACC 540
 CACCAAGTCC TTCTGTCTGC TCGTTAGCTT TAATCAATCA TGCCCTGCTT TGTCCCTCTC 600
 ACTCCCGAG CCAACCCCTA AGTGCCAAA GTGGGAGGG ACAGAGGATT CTGGGAAGCT 660
 TGAGCTCTCC CCAAGCAAT GTGAGTCCA GAGCCCGCTT TTGTCTTCC CCACAATTC 720
 ATTACTAAGA AACACATCAA ATAAACTGAC TTTTTCCTCC CAATAAAGC TCTCTTTT 780

TAATAT

Seq ID NO: 346 Protein sequence
Protein Accession #: NP_002382.1

1 11 21 31 41 51
MQHRGFLLLT LLALLALTSA VAKKDKVKK GPGSECAEW AWGPTCTPSK DCGVGPREGT 60
CGAQTQIRIC RVPCNMWKEF GADCKYKFEW WGACDGGTGT KVRQGTLLKKA RYNAQCQETI 120
RVTKFCTPKT KAKAKAKKGGK GKD

Seq ID NO: 347 DNA sequence
Nucleic Acid Accession #: NM_006783.1
Coding sequence: 1..786

1 11 21 31 41 51
ATGGATTGGG GGACGCTGCA CACTTTCATC GGGGGTGTCA ACAAACTCTC CACCAGCATC 60
GGGAAGGTGT GGATCACAGT CATCTTTATT TTCGAGTCA TGATCCTAGT GGTGGCTGCC 120
CAGGAAGTGT GGGGTGACGA GCAAGAGGAC TTCGTCTGCA ACACACTGCA ACCGGGATGC 180
AAAAATGTGT GCTATGACCA CTTTTCCTCG GTGTCCACCA TCGGCTGTG GGCCTCCAG 240
CTGATCTCTG TCTCCACCCG AGCGCTGCTG GTGGCCATGC ATGTGGCCCTA CTACAGGCAC 300
GAACCACTC GCAAGTTTCG GCGAGGAGAG AAGAGGAATG ATTCAAAGA CATAGAGGAC 360
ATTAAAAAGC ACAGGTTTCG GATAGAGGGG TCGCTGTGTT GGACGTACAC CAGCAGCATC 420
TTTTTCOGAA TCATCTTTGA AGCAGCCTTT ATGTATGTGT TTTACTTCCT TTACATGGG 480
TACCACCTGC CTTGGGTGTT GAAATGTGGG ATTGACCCCT GCGCCCAACT TGTGACTGTC 540
TTTATTTCTA GGCACACAGA GAAGACCGTG TTTACCATTT TTATGATTTC TCGCTCTGTG 600
ATTTGCTATG TCGTTAACGT GGCAGAGTTG TGCTACCTGC TGCTGAAAGT GTGTTTATAG 660
AGATCAAAGA GAGCAGACG CCAAAAAAAT CACCCCAATC ATGCCCTAAA GGAGAGTAAG 720
CAGATGAAA TGAATGAGCT GATTTTCAGT AGTGGTCAAA ATGCAATCAC AGGTTTCCCA 780
AGCTAA

Seq ID NO: 348 Protein sequence
Protein Accession #: NP_006774.1

1 11 21 31 41 51
MDWGLHIFI GGVNKHSTSI GKVNITVIFI PRVMLLVAA QEVNGDEQED FVCNTLQPGC 60
KNVCVYHFFP VSHIRLWALQ LIFVSTPALL VAMHVAYYRH ETTKFRFRGE KRNDFKDIED 120
IKKKKVRIGS SLNWTYTSST FFRITPEAAF MYVYFLYNG YHLPWVLKCG IDPCPNLVDC 180
FISRPTRTFT FTIFMISASV ICMILNVABL CYLLKVKCFR RSKRAQTQKN HPNEALKESK 240
QNEWNEIISD SQGNATIGFP S

Seq ID NO: 349 DNA sequence
Nucleic Acid Accession #: NM_002571.1
Coding sequence: 99..587

1 11 21 31 41 51
CATCCCTCTG GCTCCAGAGC TCAGAGCCAC CCACAGCCBC AGCCATGCTG TGCCTCCTGC 60
TCACCTCTGG CGTGGCCCTG GTCTGTGTTG TCCCGGCCAT GGACATCCCC CAGACCAAGC 120
AGGACCTGGA GCTCCCAAAG TTGGCAGGGA CTTGGCACTC CATGGCCATG GCGACCAACA 180
ACATCTCCCT CATGGCGACA CTGAAGCCCT CTCTGAGGCT CCACATCACC TCACTGTTGC 240
CCACCCCGA GGCACACCTG GAGATCGTTC TCACACAGATG GGAGAACAAAC AGCTGTGTTG 300
AGAGNAGGT CCTTGGAGAG AAGACTGGGA ATCCAAAGAA GTTCAAGATC AACTATACCG 360
TGGCGACAGA GGCACAGCTG CTCGATACTG ACTACAGCAA TTTCTGTGTT CTCTGCCTAC 420
AGGACCCAC CACCCCATC CAGAGCATGA TGTGCCAGTA CTTGGCCAGA GTCTGTGTTG 480
AGGACGATGA GATCATGCAG GATTCATCA GGGCTTTCAG GCGCTGCCCC AGGCACCTAT 540
GGTACTTGCT GACCTTGAAA CAGATGGGAG AGCGGTGCCG TTTCTAGCTC ACCTCGSCCT 600
CCAGGAAGAC CAGACTCCCA CCCTTCCACA CTTCCAGAGC AGTGGGACTT CTTCTGCCCC 660
TTTCAAAGAA TAACACAGCG TCAGAAAGCG ATGACGTGGT CATCTGTGTC GCCATCCCTT 720
TCCTGCTGCA CACCTGCACC ATTGCCATGG GGAGGCTGCT CCTGGGGGC AGAGTCTCTG 780
GCAGAGGTTA TTAATAAAC CTTGGAGCAT G

Seq ID NO: 350 Protein sequence
Protein Accession #: NP_002562.1

1 11 21 31 41 51
MDIPQTKQDL ELFKLAGTWH SMAMATNNIS LMATLKAPLR VHTSLDPTP EDNLEIVLHR 60
WENNSCVERK VLGETGTGPK KFKINYTVAN EATLDDTDYD NFLFLCLQDT TPIQSMHCQ 120
YLARVLVEDD EIMQGFIRAF RFLPRHLWYL LDLQKMEPC RF

Seq ID NO: 351 DNA sequence
Nucleic Acid Accession #: NM_006500.1
Coding sequence: 27..1967

1 11 21 31 41 51
ACTTGGGTCT CCGCCTCCCG CCAAGCATGG GGCCTCCAG GCTGGTCTGC GCCTTCTTGC 60
TGGCGGCTGT CTGCTGCTGT CCTCGCGTGT GGGGTGTGCC CGGAGAGGCT GAGCAGCCTG 120
CBCTTGAGCT GTTGGAGGTG GAAGTGGGCA GCACAGCCCT TCTGAAGTGC GGCCTCTCC 180
AGTCCCAAGG CAACCTCAGC CATGTGCACT GGTTTTCTGT CCACAAGGAG AAGCGGACGC 240
TCATCTTCCG TGTGCGCCAG GGCCAGGGCC AGAGCGAACC TGGGGAGTAC GAGCAGCGGC 300

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

TCAGCCTCCA GGACAGAGGG GCTACTCTGG CCTGACTCA AGTCACCCCC CAAGACGAGC 360
 GCATCTTCTT GTGCCAGGGC AAGCGCCCTC GGTCCCAGGA GTACCGCATC CAGCTCOGCG 420
 TCTACAAAGC TCCGAGAGAG CCAACCATCC AGGTCAACCC CTTGGGCATC CCTGTGAACA 480
 GTAAGGAGCC TGAGGAGGTC GCTACCTGTG TAGGGAGGAA CGGGTACCCC ATTCCCTCAAG 540
 TCATCTGGTA CAAGAATGGC CGGCCTCTGA AGGAGGAGAA GAACCGGGTC CACATTCAAT 600
 CGTCCAGAC TGTGGAGTCG AGTGGTCTGT ACACCTTGCA GAGTATTCTG AAGGCACAGC 660
 TGGTTAAAGA AGACAAAGAT GCCCAGTTT ACTGTGAGCT CAACTACCGG CTGCCAGTG 720
 GGAACCATAT GAAGGAGTCC AGGGAAGTCA CCGTCCCTGT TTTCTACCCG ACAGAAAAG 780
 TGTGCTTGA AGTGGAGCCC GTGGGAATGC TGAAGGAAGG GGACCGCTG GAATCAGGT 840
 GTTGGCTGA TGGCAACCTC CCACCACTC TCAGCATCAG CAAGCAGAAC CCCAGCACCA 900
 GGGAGGCAAG GGAAGAGACA ACCAACGACA ACGGGTCTCT GGTCTGGAG CTTGCCCGGA 960
 AGGAACACAG TGGGCTGAT GAATGTCAGG CCTGGAACCT GGACACCATG ATATCGCTGC 1020
 TGAGTGAACC ACAGGAACCT CTGGTGAAC ATGTGCTGTA CGTCCGAGTG AGTCCCGCAG 1080
 CCCCTGAGAG ACAGGAAGGC AGCAGCCTCA CCTTGACCTG TGAGGCGAG AGTAGCCAGG 1140
 ACCCTGAGTT CCAGTGGCTG AGAGAAGAGA CAGACCAAGT GCTGGAAGG GGGCCTGTGC 1200
 TTCAGTTGCA TGACCTGAAA CCGGAGGCGAG GAGGCGGCTA TCGCTGCGTG GCGTCTGTGC 1260
 CCAGCATACC GCGCCTGAAC GCACACAGC TGGTCAAGCT GGCCATTTT GGCCCCCTT 1320
 GGATGGCATT CAAGGAGAGG AAGGTGTGGG TGAAAGAGAA TATGGTGTG AATCTGTCTT 1380
 GTGAAGCGTC AGGGCACCCC CGGCCACCA TCTCTGGAA CGTCAACGCG ACGGCAAGTG 1440
 AACAGAGACA AGATCCACAG CGAGTCTCTG GCACCTGAA TGTCTCTGT ACCCGGAGC 1500
 TGTTCGAGAC AGGTGTTGAA TGCAOCCCTT CCAACGACCT GGGCAAAAC ACCAGCATCC 1560
 TCTTCTGCA GCTGGCTAAT TTAACACCC TCACACAGA TCCCAACACA ACCACTGGCC 1620
 TCAGCATTTC CACTGCCAGT CCTCATACCA GAGCCACAG CACCTCCACA GAGAGAAAGC 1680
 TGGCGGAGCC GGAGAGCCGG GCGTGGTCA TCGTGGCTGT GATTGTGTGC ATCTGGTTC 1740
 TGGCGGTGCT GCGCCTCTGC CTCTATTTC TCTATAAGAA GGGCAAGCTG CCGTGCAGGC 1800
 GCTCAGGGA GCAGGAGATC ACGCTGCCCT CGTCTCGTAA GACCGAAGCT GTAGTTGAAG 1860
 TTAAGTCAGA TTAGCTCCCA GAAGAGATGG GCTCTCTGCA GGGCAGCAGC GGTGACAGA 1920
 GGGCTCCGG AGACGAGGA GAGAAATACA TCGATCTGAG GCATTAGCCC GGAATCACTT 1980
 CAGCTCCCTT CCTGCCCTGG ACCATTCCCA GCTCCCTGCT CACTCTCTC TCAGCCAAAG 2040
 CCTCCAAAG GACTAGAGAG AAGCCTCTCT CTCCCTCAC CTGACACCC CCTTTCAGAG 2100
 GGGCAGTGGG TTAGGACCTG AGGACCTCAC TTGGCCTGCG AAGCGCTTT TCAGGGAACA 2160
 GTCCACACAG ATCTCTCTCA CTTGAGTGA AGCTCATCCC AAGCAAGGAG CCCAGTCTC 2220
 CCGAGCGGGT AGGAGAGTTT CTGTCAGAAC GTGTTTTTC TTACACACA TTATGGCTGT 2280
 AAATACCTGG CTCTGCGCAG CAGCTGAGCT GGGTAGCCTC TCTGAGCTCG TTTCTGCCC 2340
 CAAAGGCTGG CTTCACCAT CCAGGTGCAC CACTGAAGTG AGGACACCC GAGGCCAGGC 2400
 GCGTCTCAT GTTGAAGTGC GCTGTTCACA CCGCTCCGG AGAGCACCCC AGCGCATCC 2460
 AGAAGCAGCT CAGTGTGTGC TGCCACCACC CTCCGCTCG CCTCTTCAA GTCTCTGTG 2520
 ACATTTTTC TTGTTTCAGA AGCCAGGAAC TGGTGTCTAT CCTTAAAGA TAOGTGGCG 2580
 GGGCAGGTGT GGTGGCTCAC GCCTGTAAAT CCAGCACTTT GGGAGGCGGA GGGGGCGGA 2640
 TCACAAAGTC AGGACGAGAC CATCTGGCT AACACGGTGA AACCTGTCT CTAATAAAA 2700
 TACAAAAAA AATTAGCTAG GCGTAGTGT TGGCACCTAT AGTCCAGCT ACTCGAAGG 2760
 CTGAGCAGG AGAATCGTAT GAATCCAGGA GGTGGAGCTT GCAGTGAGCC GAGACCGTGC 2820
 CACGCACTC CAGCCTGGGC AACACAGGGA GACTCCGTCT CGAGGAAAAA AAAAGAAAA 2880
 AGCGTAGCT GCGGTGAGGA AGCTGGGCGC TGTTTTCGAG TTCAGTGAA TTAGCTCAA 2940
 TCCCGTGT CTCTGCTCC CATAGCCTC TTGATGATC AOSTAAACT GAAAGGCAGC 3000
 GGGGAGCAGA CAAGATGAG GTCTACACTG TCCCTCATGG GGAATTAAGC TATGGTTATA 3060
 TTAGACCAA ACTTCTACAA ACCAAGCTCA GGGCCCAAC OCTAGAAGGG CCAAAATGAG 3120
 AGAATGGTAC TTAGGATGAG AAACGGGGC CTGGCTAGAG CTTCGGGTGT GTGTGTCTGT 3180
 CTGTGTAT GCATACATAT GTGTGTATAT ATGTTTTGT CAGGTGTGTA AATTGCAAA 3240
 TTGTTTCTT TATATATGTA TGTATATATA TATATGAAA TATATATATA TAIGAAAAAT 3300
 AAAGCTTAT TGTCCCAAG AATCATACAT TGCTTTTTA TTCTACATGG GTACCCAGG 3360
 AAACCTGGGG CCTGTGAAAC TACACACAAA AGGCACACAA AACCGTTCC AGTTGGCAGC 3420
 AGAGATCAGG GGTATCTCT GCTTCTGAGC AAATGGCTCA AGCTTACCA GAGCAGACAG 3480
 CTACCTACT TTTGAGCAGC AAAACGTCCC GTATGACGCA GCACGAAGGG CTTGGCAGGC 3540
 TGTAGCAGG AGCTATGTCC CTTCCTATCG TTTCGTGCA CTT

Seq ID NO: 352 Protein sequence
 Protein Accession #: NP_006491.1

60
 65
 70

1 11 21 31 41 51
 GLFRLVCAFL LAACCCCPRV AGVPGABQP APELVEVEVG STALLKCLSL QSQGNLSHVD 60
 WFSVEKEKRT LFRVROGG QSEPGYEQR LSLQDRGATL ALTQVTPQDE RIFLCQGRKP 120
 RSQEYRIQLR VYKAFEPENI QVNPLGIPVN SKEPEEVATC VGRNGYPIBQ VIWYKGRFL 180
 KEKENVHIQ SSQTVESGGL YTLQSLKQAQ LVKEDKDAQF YCEINYLIFS GNMHKESEV 240
 TVPVFPTPEK VNLVEVPVGM LKESDRVEIR CLADGNPPPH PSISQNPST REABETIND 300
 NGVLVLEPAR KEHSGRYECQ ANNLDTMISL LSEFQELLVN YVSDVRVSPA APERQEGSSL 360
 TLTCRAESSQ DLEFQMLREE TDQVLERGPV LQLHDLKREA GGYRCVASV PSIFGLNETQ 420
 LVKLAIFGPF WMAFKERKVV VKENMVLNLS CEASGHPRPT ISNNVNGTAS EQDQDPQVRL 480
 STLNLVLTPE LLETGVECTA SMDLGNKTSI LFLNLVNLIT LTPDSNTTGT LSTSTASPHY 540
 RANSTSTERK LPPEPSRGVV IVAVIVCLLV LAVLGAVLYF LYKKGKLPKR RSGRQKITLP 600
 PSRTELVEE VKEDKLPEEM GLQSSGDK RAPGDQGERY IDLRL

75
 Seq ID NO: 353 DNA sequence
 Nucleic Acid Accession #: NM_003183.3
 Coding sequence: 165..2639

80

1 11 21 31 41 51
 TCGAGCCTGG CGGTAGAATC TTCCCACTAG GCGGCGCGGG AGGCAAAAGA GGATTGAGGG 60
 GCTAGGCCGG GCGGATCCCG TCTCCCTCCG ATGTGAGCAG TTTTCCGAAA CCCCGTCAGG 120
 CGAAGGCTGC CCAGAGAGGT GGAGTCGGTA GCGGGCGCGG GAACATGAGG CAGTCTCTCC 180
 TATTCTGAC CAGCGTGGTT CCTTCTGTGC TGGCGCGCGG ACCTCCGAT GACCGGGBCT 240
 TCGCCCCCA CCAGAGACTC GAGAAGCTTG ATTCTTTGCT CTCAGACTAC GATATTCTCT 300

CTTTATCTAA TATCCAGCAG CATTCCGTAA GAAAAAGAGA TCTACAGACT TCAACACATG 360
 TAGAAACACT ACTAACITTT TCAGCTTTGA AAAGGCATTT TAAATTATAC CTGACATCAA 420
 GTACTGAACG TTTTTCACAA AATTTCAAGG TCGTGGTGGT GGATGGTAAA AACGAAAGCG 480
 AGTACACTGC AAAATGGCAG GACTTCTTCA CTGGACACGT GGTGGGTGAG CCTGACTCTA 540
 GGGTTCTAGC CCACATAAGA GATGATGATG TTATAATCAG AATCAACACA GATGGGGCCG 600
 AATATAACAT AGAGCCACTT TGGAGATTGG TTAATGATAC CAAAGACAAA AGAATGTTAG 660
 TTTATAAATC TGAAGATATC AAGAATGTTT CAGCTTTGCA GTCTCCAAAA GTGTGTGGTT 720
 AITTAAGAGT GGATAATGAA GAGTTGCTCC CAAAAGGGTT AGTAGACAGA GAACCACTGT 780
 AAGAGCTTGT TCATCGAGTG AAAAGAAAGAG CTGACCCAGA TCCCATGAAG AACACGTGTA 840
 AATTATTGGT TCGAGCAGAT CATCGCTTCT ACAGATACAT GGGCAGAGGG GAAGAGAGTA 900
 CAACTACAAA TTACTTAATA GAGCTAATTG ACAGAGTTGA TGACATCTAT CGGAACACTT 960
 CATGGGATTA TGCAGGTTT AAAGGCTATG GAATACAGAT AGAGCAGATT CGCATCTCA 1020
 AGTCTCCACA AGAGGTAAAA CCTGGTGAAA AGCACTACAA CATGGCRAAA AGTTACCCAA 1080
 ATGAAGAAAA GGATGCTTGG GATGTGAAGA TGTGTCTAGA GCAATTTAGC TTTGATATAG 1140
 CTGAGGAAGC ATCTAAAGTT TGCCTGGCAC ACCTTTTCAC ATACCAAGAT TTTGATATAG 1200
 GAACTCTTGG ATTAGCTTAT GTTGGCTCTC CCAGAGCAAA CAGCCATGGA GTGTGTTGTC 1260
 CAAAGGCTTA TTTGAGCCCA GTTGGGAAGA AAAATATCTA TTTGAATAGT GGTGTGACGA 1320
 GCACAAAGAA TTATGTTAAA ACCATCCCTA CAAAGGAAGC TGACCTGGTT ACAAATCATG 1380
 AATTGGGACA TAATTTTGA GAGAAACATG ATCCGGATGG TCTAGCAGAA TGTGCCCCGA 1440
 ATGAGGACCA GGGAGGGAAA TATGTCTATG ATCCCATAGC TGTGAGTGGC GATCAGAGTA 1500
 ACAATAGAT GTTTTCAAA TGCAGTAAAC AATCAATCTA TAAGACCATT GAAAGTAAGG 1560
 CCAGGAGATG TTTTCAGAAA GGCAGCAATA AAGTTTGTGG GAACTTCAGG GTGGATGAAG 1620
 GAGAAGAGTG TGATCTTGGC ATCATGTATC TGAACCAAGA CACCTGCTGC AACAGCGACT 1680
 GCACGTGAAA GGAAGGTGTC CAGTGCAGTG ACAGGAACAG TCCTTGCTGT AAAAATCTGT 1740
 AGTTTGAGAT GGTCCAGAG AAGTCCAGG AGGCGATTAA TGCTACTTGC AAAGGCTGTG 1800
 CCTACTCGCT AGGTAAATAG AGTGAAGTGC CGCTCCAGG AATGCTGAA AATGACTCTG 1860
 TTTGCTTGGC TTTTGGCAAG TGTAAAGATG GGAATGTCAT CCCTTTCTGC GAGAGGGAAAC 1920
 AGCAGCTGGA GTCTGTGCA TGTAAATGAAA CTGACAACTC CTGCAAGGTG TGTGCGAGGG 1980
 ACCCTTCTGG CCGCTGTGTG CCCTATGTGG ATGCTGAACA AAAGAACTTA TTTTGAAGTA 2040
 AAGGAAGGCC CTGTACAGTA GGAATTTGTG ACATGAATGG CAAATGTGAG AAACGAGTAC 2100
 AGGATGTAAT TGAACGATTT TGGGATTTC TTAGACAGCT GAGCATCAAT ACTTTTGGAA 2160
 AGTTTATAGC AGACCAATG GTTGGGTCTG TCCTGTGTTT CTCTTGATA TTTTGGATTG 2220
 CTTTACAGAT TCTTGTCCAT TGTGTGGATA AGAATTTGGA TAACAGTAT GAACTCTCTG 2280
 CTCTGTTTCA CCGCAGTAA GTCGAAATGC TGAGCAGCAT GGATTTCTGA TCGGTTCTGA 2340
 TTATCAAAAC CTCTCTGCG CCCCAGACTC CAGGCCGCTT CGAGCCTGCC CCTGTGATCC 2400
 CTCTGGGCCC AGCAGCTCCA AAACCTGGAC ACCAGAGAAAT GGACACCATC CAGGAAGACC 2460
 CCAAGCAGTA CTCCAGTAA GACGAGGATG GGTGTGAGAA GGACCCCTTC CCAATATAGA 2520
 GCACAGCTGC CAAGTCAATT GAGGATCTCA CGGACCATCC GGTCCGCAAG AGTGAAAGG 2580
 CTGCTCTCTT TAACTGCGAG CGTCAGAAAT GTGTTAACAG CAAAGAAACA GAGTGCTAAT 2640
 TTAGTCTTCA CTCTTCTGA CTAAAGTGTG CAAATATTTT TTATAGATT GACCTACAAA 2700
 TCAATCAGAG CTGTATTTT GTGAAGACTG GGAAGTGAAT TAGCAGATGC TGTGATGTG 2760
 TTTGAACCTC CTCCAGTAA ACAGTCTCTG TGTGTTTGG CCCTCTCTCT TTTGAAAGG 2820
 TAAGGTGAAA GTGAATCTAC TTATTTTGG GCTTTCAGGT TTTAGTTTT AAAATATCTT 2880
 TTGACCTGTG GTGCAAAAGC AGAAAATACA GCTGGATTGG GTTATGAATA TTTAGTTTT 2940
 TGTAAATTA TCTTTTATAT TGATAACAGC ACTGACTAGG GAAATGATCA GTTTTTTTT 3000
 ATACACTGTA ATGACCGCTT GAATATGAAG CATTGGGCAT TTATTTGTGA GAAAGTGA 3060
 ATAGTTTTT TTTTTTTTT TTTTTTTTG CTCAACTAA AAACAAAGGA GATAAATTA 3120
 GTATACATTG TATCTAAAT GTGGGCTAT TCTAGTTAT TACCCAGAGT TTTATGTAG 3180
 CAGGGAATAT ATATATCTAA ATTTAGAAAT CATTGGGTT AATATGGCTC TTCTAATTC 3240
 TAAAGTAAT GCTCAGAAAC TAACCACTAC CTACAGTGA GGGCTATACA TGTAGCCAG 3300
 TTGAATTTAT GGAATCTACC AACTGTTTAG GGCCCTGATT TGCTGGGAGG TTTTCTGTA 3360
 TTTTATAGT ATCTTCATGT ATCCCTGTTA CTGATAGGGA TACATGTCTT AGAAAATTA 3420
 CTATTGGCTG GGAATGTTGG CTCATGCTGT TAATCCAGC ACTTGGAGAG GCTGAGGTTG 3480
 CGCCACTACA CTCCAGCTG GGTGACAGAG TGAGATCTGC CTC

Seq ID NO: 354 Protein sequence
 Protein Accession #: NP_003174.2

1 11 21 31 41 51
 MRQSLLFLTS VVPFLAFRP EDDPGFGPHQ RLEKLSLLS DYDILSLSNL QHSEVRKRDL 60
 QTSTHETLL TFSALKRHFK LYLTSSTERF SQNFKVVVD GKNSEYTAQ WQDFFTHVV 120
 GEPDSRVLAH IRDDIVIRI NTDGAEYNIE PLNRFPVNDK DKRMILVYKSE DIKRVSLQS 180
 FKVCGLKVD NEELLPKGLV DREPFEEVLH RVKRRADPDF MKNTCKLNV ADHREYRYMG 240
 RGEESTPTNY LIELIDRVDD IYRNTSWDMA GFKGYGQIE QIRILKSPQE VKPGEKHYNM 300
 AKSYENNEKD ANDVKMLLEQ FSPDIABEAS KVCLAHFTY QDFDMGTLGL AYVGSPPRNS 360
 HGVCPKAYY SPVGMKIYH NSGLTSTKNY KRTILTKBAD LVTTHLGRN FGAEEDPDGL 420
 AECAPNEDQG GKYVMYPIAV SGDHENKMF SNCSKQSIYK TIESKAQECF QERENKVCWN 480
 SRVDEGEBCD PGIMYLNNDT CCNSDCTLKE GVQCSDRNSP CCKNCQFETA QKKQQAANA 540
 TCKGVSYCTG NSSECPFFGN ARNDTVCLDL GKCDGKICP FCEREQQLS CACNETDNC 600
 KVCCRDLSGR CUPYVDABQK NLPLRKGKPC TVGFCDMNGK CEKRVQDVIE RFNDFIDQLS 660
 INTFGKILFD NIVGSLVFS LIFWIPFSIL VECVDKLDK QYESLSLPHF SNVEMLSMD 720
 SASVRIKPF PAFQTPGRILQ PAFVIPSAPA APKLDHQRMD TIQEDPSTDS HMDGDFEED 780
 PFFNSSTAAR SFEDLIDHPV ARSEKAASEK LQRQNRVNSK ETEC

Seq ID NO: 355 DNA sequence
 Nucleic Acid Accession #: NM_021832.1
 Coding sequence: 164..2248

1 11 21 31 41 51
 TCGAGCTTGG CGGTAGAATC TTCCAGTAG GCGGCGCGGG AGSAAAAGAG GATTGAGGGG 60
 CTAGGCCGGG CGGATCCCGT CCTCCCCCGA TGTGAGCAGT TTTCCGAAAC CCCGTCAGGC 120
 GAAGGCTGCC CAGAGAGGTG GAGTCGATG CGGGCCCGGG AACATGAGGC AGTCTCTCCT 180

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

ATTCTGTACC AGCGTGGTTC CTTCGTGCT GCGCGCCGCA CCTCCGGATG ACCCGGGCTT 240
 CGGCCCCAC CAGAGCTCG AGAAGCTTGA TTCTTGTCTC TCAGACTACG ATATTCTCTC 300
 TTATCTAAT ATCCAGCAGC ATTCGGTAAAG AAAAAGAGAT CTACAGACTT CAACACATGT 360
 AGAAACACTA CTAACCTTTT CAGCTTTGAA AAGGCATTTT AAAATTATACC TGACATCAAG 420
 TACTGAACGT TTTTCACAAA ATTTCAAGGT CGTGGTGGTG GATGGTAAAA ACGAAAGCGA 480
 GTACACTGTA AAATGGCAGG ACTTCITCAC TGGACACGTG GTTGGTGAGC CTGACTCTAG 540
 GGTTCAGCC CACATAAGAG ATGATGATGT TATAATCAGA ATCAACACAG ATGGGGCCGA 600
 ATATAACATA GAGCCACTTT GGAGATTGT TAATGATACC AAAGACAAA GAATGTTAGT 660
 TTATAAATCT GAAGATATCA AGAATGTTTC ACGTTTGCAG TCTCCAAAAG TGTGTGGTTA 720
 TTTAAAGTG GATAATGAAG AGTTGCTCCC AAAAGGGTTA GTAGACAGAG AACCACCTGA 780
 AGAGCTTGT CATCGAGTGA AAAGAAGAGC TGACCCAGAT CCCATGAAGA ACAGGTGTAA 840
 ATTATTGGTG GTAGCAGATC ATCGCTTCTA CAGATACATG GGCAGAGGGG AAGAGAGTAC 900
 AACTACAAAT TACTTAATAG AGCTAATTGA CAGATTTGAT GACATCTATC GGAACACTTC 960
 ATGGGATAAT GCAGGTTTGA AAGGCTATGG AATACAGATA GAGCAGATTG GCATCTCAA 1020
 GTCTCCACAA GAGGTAAAC CTGGTGAAAA GCACTACAAC ATGGCAAAA GTTACCCAAA 1080
 TGAGRAAAG GATGCTTGGG ATGTGAAGAT GTTGTAGAG CAATTAGCT TTGATATAGC 1140
 TGAGGAAGCA TCTAAAGTTT GCTTGGCACA CCTTTTCA TAACCAAGAT TTGATATGGG 1200
 AACTCTTGA TTAGCTTATG TTGGCTCTCC CAGAGCAAC AGCCATGGAG GTGTTTGTCC 1260
 AAAGGCTAT TTTCGCAAG TTGGGAAGAA AATATCTAT TTGAATAGTG GTTTGACGAG 1320
 CACAAAGAAT TATGTAATA CCATCCITAC AAAGGAAGCT GACCTGGTTA CAATCATGA 1380
 ATTGGGACAT AATTTTGGAG CAGAACATGA TCCGGATGGT CTAGCAGAT GTGCCCGGAA 1440
 TGAGGACCC TGAGGGAAT ATGTCATGTA TCCCATAGCT GTGAGTGGCG ATCAGAGAA 1500
 CAATAAGATG TTTTCAACT GCAGTAAACA ATCAATCTAT AAGACCATG AAAGTAAGGC 1560
 CCAGGAGTGT TTTCGCAAG GTAGCAATAA AGTTTGTGGG AACTCGAGGG TGGATGAAGG 1620
 AGAAGAGTGT GATCTGGCA TCATGATCT GAACAACGAC ACCTGCTGCA ACAGCGACTG 1680
 CAGTTGAAG GAAGGTGTC AGTGCAAGT CAGGAACAGT CCTGTCTGTA AAAACTGTCA 1740
 GTTTGAGACT GCCCAGAAGA AGTGCCAGGA GCGGATTAAT GCTACTTGCA AAGGOGTGTG 1800
 CTACTGCACA GGTAAATAGA GTGAGTGCCG GCCTCCAGGA AATGCTGAAG ATGACACTGT 1860
 TTGCTTGGAT CTTCGCAAG GTAAGGATGG GAAATGCATC CCTTTCIGCG AGAGGGAACA 1920
 GCAGCTGGAG TCCGTGTCAT GTAATGAAC TGACAACCTC TGCAAGGTGT GCTGCGAGGA 1980
 CCTTCCGCG CGCTGTGTGC CCTATGTGTA TGTGAACAA AGAACITAT TTTTGAGGAA 2040
 AGGAAGCCG TGTACAGTAG GATTTGTGA CATGAATGGC AAATGTGAGA AACGATACA 2100
 GGAATGTAAT GAACGATTTT GGAATTTTAT TGACAGCTG AGCATCAATA CTTTGGGAAA 2160
 GTTTTATGTA GACACATCG TTGGTCTGT CCTGGTTTC TCCTTGATAT TTTGGAITCC 2220
 TTTTCACTT CTGTCTCAT GTGTGTAAC TGGAATGCT GAGCAGCATG GATTCTGCAT 2280
 CGGTTCGAT TATCAAAACC TTCTCTGCG CCCAGACTCC AGGCCGCTG CAGCTGCCC 2340
 CTGTGATCCC TTGCGGCGCA GCAGCTCCAA AACTGAGCCA CCAGAGAATG GACACCATCC 2400
 AGGAAGACCC CAGCAGAGC TCACATATGG ACGAGGATGG GTTTGAGAAG GACCCCTTCC 2460
 CAAATAGTAG CACAGCTGCC AAGTCATTTG AGGATCTCAC GGACCATCCG GTCACCAAAA 2520
 GTGAAAGGCG TGCTCTCTT AACTGCAGC GTCAGAACTG TGTGACAGC AAAGAAACAG 2580
 AGTGCTAAT TAGTCTCAG CTCTCTGAC TTAAGTGTG AAAATATTTT TATAGATTTG 2640
 ACCTACAAAT AATCAGACT TATATTTGT GAAGACTGGG AAGTGACTTA CGAGTGTCTG 2700
 GTCTGTGTT TGAATCTCT GCAGTAAAC AGTTCTGTG TGGTTTGGCC CTCTCTCTT 2760
 TGAAAGGTA AGTGAAGGT GAATCTAGCT TATTTTGGG CTTCAGGTT TTATGTTTGA 2820
 AAATATCTTT TGACCTGTGG TGCAAAAGCA GAAATACAG CTGGATTGGG TTATGAGTAT 2880
 TTAGCTTTT GTAAATTAAT CTTTATAT TATAACAGGC ACTGACTAGG GAAATGATCA 2940
 GTTTTTTT ATACACTGTA ATGAACGCT GAATATGAAG CATTTGGCAT TTATTGTGTA 3000
 GAAAGTGA ATAGTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 3060
 GATAAATTA GTATCAATG TATCTAAAT GTGGGTCTAT TTCTAGTTAT TACCAGAGT 3120
 TTTTATGTAG CAGGAAAT ATATATCTAA ATTTAGAAAT CATTTGGGTT AATATGGCTC 3180
 TTCATAATC TAAGACTAAT GCTCAGAAC TAACCACTAC CTTACAGTGA GGGCTATACA 3240
 TGGTAGCCAG TTGAATTTAT GGAATCTACC AACTGTTAG GGGCTGATT TGCTGGGCG 3300
 TTTTCTGTA TTTTATAAGT ATCTTCAAT ATCCCTGTTA CTGATAGGGA TACATGTCTT 3360
 AGAAATTA CTATTTGCTG GAGTGGTGG CTCATGCCG TAAATCCAGC ACTTGGAGAG 3420

Seq ID NO: 356 Protein sequence
 Protein Accession #: NP_068604.1

60
 65
 70

1 11 21 31 41 51
 MRQSLLELTS VVPFVLAPRP PDDEPGFGPHQ RLEKLDLLS DYDILSLSNL QHSHVRKRD 60
 QTSTHVELL TFSALKRHPK LYLTSSTERF SQNFKVVVVD GENSEBYTVK WQDFFTGHVV 120
 GEPDSRVLAH IRDDOVIIRI NTDGAHYNLE PLNRFFVMDTK DKRMLVYKSE DIKNVSRLOS 180
 PKVCGYLKVD NEELLPKGLV DREPPPELVE RVKRRADDPD MKNTCKLLVV ADHRFYRYMG 240
 RGEESTTNY LIEHLIDRVD IYRNTSDWNA GFKGYGQIE QIRILKSPQE VKPGEKHYNM 300
 AKSYFNEEKD ANDVRMLLEQ FSPDIAEAS KVCLAHIFTY QDFDNGTLGL AYVGSPRANS 360
 HGVCPKAYY SPVGEKNLYL NSGLTSTENY GKTILTKEAD LVTTHLGNH FGAEHPDGL 420
 AECAPNEDQG GKYVMPIAV SGDHENNKMF SNCSKQSIYK TIESKQRCF QERSNKVCGN 480
 SRVDEGRECD PGIMVLMNDT CQNDCTLKE GVQCSDRNSP CCKNCQFETA QKKQCBALNA 540
 TCKGVSYCTG NSSECPFPNG AEDDTVCLDL GKCDGKICP FCERREQLES CACNETDNSC 600
 RVCCRDLSGR CVFVVDABQK NLFRLKSKPC TVGFCDMNGK CCKRVQDVIE RFWDFIDQLS 660
 INTFGKFLAD NTVGSVLVFS LIFWIPFSIL VHCV

Seq ID NO: 357 DNA sequence
 Nucleic Acid Accession #: NM_004994.1
 Coding sequence: 20..2143

80

1 11 21 31 41 51
 AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCTG GTCTGTGTCT TCCTGGTGTCT 60
 GGGCTGCTGC TTGCTGCCCC CCAGACAGCG CCAGTCCACC CTTGTGCTCT TCCTGGGAGA 120
 CCTGAGAAC AATCTCACCG ACAGGAGCT GGCAGAGGAA TACCTGTACC GCTATGTTA 180
 CACTCGGGTG GCAGAGATGC GTGGAGATC GAAATCTCTG GGGCTGCGC TGCTGCTTCT 240

5
 10
 15
 20
 25
 30
 35

CCAGAAAGCAA CTGTCCCTGC CCGAGACCBB TGAGCTGGAT AGCGCCACGC TGAAGGCCAT 300
 GCGAAACCCCA CGGTGCGGGG TCCAGACCT GGCAGATTTC CAAACCTITG AGGGGACCTT 360
 CAAGTGGCACC CACCACAACTA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTGCCTGG 420
 GCGGTGATTT GACGACGGCT TTGCCGCGGC CTTCGCACTG TGGAGCGCGG TGACGCCGCT 480
 CACCTTCACT CGCGTGTACA GCGGGGACGC AGACATCGTC ATCCAGTTTG GTGTGCGGGA 540
 GCACGGAGAC GGGTATCCCT TCGACGGGAA GACGGGCTC CTGGCACACG CCTTTCCTCC 600
 TGGCCCGGGC ATTCAGGGAG ACBCCCATTT CGACGATGAC GAGTTGTGTT CCTTGGGCAA 660
 GGGCGTGGTG GTTCAACTTC GGTITGGAAA CGCAGATGGC GCGGCTTGGC ACTTCCCTCT 720
 CATCTTCGAG GGCCTCTCT ACTCTGCTTG CACACCGAC GGTGCTCCG ACCGCTTGCC 780
 CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTGGGTCTCT GCGCCAGCGA 840
 GAGACTCTAC ACCCGGGACG GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCACTTT 900
 CCAAGGCCAA TCTTACTCG CCTGCACAC GACCGGTGGC TCGACGGCT ACCGCTGGTG 960
 CGCCACACAC GCGAACTAGC ACCGGGACAA GCTCTTCGGC TTCGCGCGA CCGAGCTGA 1020
 CTGACGGTG ATGGGGGGCA ACTCGCGGG GAGCTGTGC GTCTTCCCT TCACTTCTCT 1080
 GGGTAAGGAG TACTCGACCT GTACGACBA GGGCGCGGGA GATGGGCGCC TCTGGTGGC 1140
 TACCACCTCG AACCTTGTACA GCGACAGAA GTGGGGCTTC TGCCCGGACC AAGGATACAG 1200
 TTTGTCTCTC GTGGCGGGC ATGAGTTGCG CCAAGCGCTG GGTCTAGATC ATTCTCAGT 1260
 CGCGGAGGCG CTATGTATCC CTATGTACCG CTTCACGTAG GGGCCCCCT TGCATAAGGA 1320
 CGACGTGAT GCGATCGGGC ACCTCTATGG TCTCTGCTCT GAACCTGAGC CACGGCTCTC 1380
 AACCACACAC ACACCGCAGC CCACGGCTCC CCCGACGGTC TGCCCCACCG GACCCCCCAC 1440
 TGCTCACCCC TCAGAGCGCC CCACAGCTGG CCCACAGGT CCCCCTCAG CTGGCCCCAC 1500
 AGGTCCCGCC ACTGCTGGCC CTCTACGGC CACTACTGTG CCTTTGAGTC CGGTGGACGA 1560
 TGCTGTCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAAACGAG TGTATTGTGT 1620
 CAAGGATGGG ARGGTCTGGC GATTCTCTGA GGGCAGGGG AGCGGGCGCG AGGGCCCCCT 1680
 CTTTATGCGC GACAACTGGC CGCGCTGCCC CGGCAAGCTG GACTCGGTCT TTGAGGAGCC 1740
 GCTCTCCAGG AACCTTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGGCTC 1800
 GGTGCTGGCG CGGAGCGCTC TGGACAAGCT GGGCTGGGA GCGACGCTGG CCCAGGTGAC 1860
 GGGGGCCCTC CGGAGTGGCA GGGGGAAGAT GCTGCTGTT AGCGGGCGCG GCTCTGAGG 1920
 GTTCGACGTG AAGGCGCAGA TGGTGGATCC CCGGAGCGCC AGCGAGGTGG ACCGGATGTT 1980
 CCGCGGGTG CCTTTGGACA CGCACGACHT CTTCCAGTAC CGAGAGAAAG CCTATTCTGT 2040
 CCAGGACCGC TTTCTACTGC GCGTGAATTC CCGGAGTGG TTGAACCAAG TGCACCAAGT 2100
 GGGCTACGTG ACCATATGACA TCTGCAAGTG CCTGAGGAC TAGGGCTCCC GTCTGCTTT 2160
 GCACTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGGA AGGAGCCAGT TTGCGCGATA 2220
 CAAACTGGTA TTTCTGTTCT GAGGAAAGGG AGGAGTGGAG GTGGGCTGGG CCTCTCTCTC 2280
 TCACCTTTGT TTTTGTGTG AGTGTCTCTA ATAACTTGG ATTCTCTAAC CTTT

Seq ID NO: 358 Protein sequence
 Protein Accession #: NP_004985.1

40
 45
 50

1 11 21 31 41 51
 MSIMQPLVLV LLLVLCPCAA PRQRQSTLVL FPGDLRTNLT DRQLAEYLY RYGYTRVAEM 60
 RGSKSLGPA LLLQLQQLSL PETGELDSAT LKAMRTPRCG VPDLGRFQTF EGDLEWHHHN 120
 ITYNIQNYSE DLPRAVIDDA FARAFALWSA VTPLTFTVRY SRDADIVIQF GVAERGDGYP 180
 FDGKGLLAH AFPPGPGIQG DAHFDDDELW SLGKGVVVPF RFGNADGAC HFFFIQEGRS 240
 YSACTTDGSR DGLPWCSTTA NYDIDDFRFG CPSERLYTRD GNADGKPCQF PFIFQQQSYS 300
 ACTIDRGRDG YRNCATTANY DRDLFGFCFP TRADSTVNGG NSAGELCVFP PFLGKEYST 360
 CTSRGRGDGR LMCATITNFD SDKINGFCPD QGYSLFLVAA HEFGHALGLD HSSVPEALMY 420
 EMYRFTGEGP LHKDDVNGIR HLYGPRFEPE PRPPTTTTTF PIAPPTVTCPT GPFTVHPSE 480
 PTAGTGPFPF AGPTGPPTAG PSTATTVPLS FVIDRNVNI FDLAEIGNQ LYLFKDKKYW 540
 RFSRGRSRP GQFPLADKN PALPRKLDVS FKEPLSKLF FFSGRQVWVY TGSVLGPRR 600
 LDKLGLADV AQTGALRSR RGRMLFSGR RLWRFDVKAQ MVDPRSABEV DRMFPGVPLD 660
 TRDVFQYREK AYFCQDRPYW RVSSRSRLNQ VDQVGVVTD ILQCPED

55
 Seq ID NO: 359 DNA sequence
 Nucleic Acid Accession #: NM_000213.1
 Coding sequence: 127..5385

60
 65
 70
 75
 80

1 11 21 31 41 51
 CGCCCGCBG CTGCAGCCCC ATCTCCTAGC GGCAGCCGAG GCGCGGAGGG AGCGAGTCCG 60
 CCGCGAGGTA GGTCCAGGAC GCGCGCACAG CAGCAGCCGA GGTGCGCGG GAGAGGGAGG 120
 AAGAGGATGG CAGGGCCACG CCGCAGCCCA TGGGCCAGGC TGCTCCTGGC AGCCTTGATC 180
 AGGTCAGCG TCTCTGGGAC CTCTGCAAAC CGCTGCAAGA AGGCCCCAGT GAAGAGCTGC 240
 ACGGAGGTG TCGGTGTGGA TAAGGACTGC GCTTACTGCA CAGACGAGAT GTTCAGGGAC 300
 CGCGCTGCA ACACCCAGGC GAGAGTGTCT GCGCGGGCT GCGAGCGGA GAGCATCGTG 360
 GTCATGGAGA GCAGCTTCCA AATCACAGAG GAGACCCAGA TTGACACCAC CCGCGGCGC 420
 AGCCAGATGT CCGCCCAAGS CCGTGGGGTC GCTCTGGGCG CGGTGAGGA GCGGCATTTT 480
 GAGCTGGAGG TCTTTGAGCC ACTGGAGAGC CCGTGGGACC TGTACATCCT CATGGACTTC 540
 TCCAATCCA TGTCCATGA TCTGGACAAC CTCAGAAGA TGGGGCAGAA CCGGCTCGG 600
 GTCTGTGACC AGCTCACCAG CAGTACACT ATTGGATTG GCAAGTTGT GCACAAATC 660
 AGGTCCTCG AGACCGACAT GAGGCTGAG AAGCTGAAG AGCCCTGGCC CAACAGTGAC 720
 CCGCCCTCT CTTTCAAGAA CGTCAACAGC CTGACAGAAG ATGTGATGA GTTCCGGAAT 780
 AAACCTGAG GAGAGCGGAT CTCAGGCAAC CTGGATGCTC CAGAGGGCGG CTTGATGCC 840
 ATCTGCAGA CAGCTGTGTG CACGAGGGAC ATTGGCTGGC GCGCGGACAG CACCCACCTG 900
 CTGGTCTTCT CCGCCAGTGC AGCCTTCCAC TATGAGGCTG ATGGCGCAA CBTGCTGGCT 960
 GGCATCATGA GCGCGAACGA TGAACGGTGC CACTGGACA CCACGGGAC CTACACCCAG 1020
 TACAGGACAC AGGACTACCC GTGGTGGCCC ACCCTGGTGC GCTGCTGCG CAGACACAC 1080
 ATCATCCGTA TCTTGTCTGT CACCAACTAC TCTATAGCT ACTACAGAA GCTTCAACAC 1140
 TATTTCCCTG TCTCTCACT GGGGTGTCTG CAGGAGGACT CGTCCACAT CGTGGAGCTG 1200
 CTGGAGGAGG CCTTCARTCG GATCGGCTCC AACTTGGACA TCGGGCCCT AGACAGGCC 1260
 CGAGCCCTTC GACACAGAGT CACCTCCAG ATGTTCCAGA AGACGAGGAC TGGGTCTCTT 1320
 CACATCCGCG GGGGGGAGT GGTATATAC CAGGTGCAGC TCGGGGCCCT TGAGCACCTG 1380
 GATGGGACGC ACGTGTGCCA GCTGCCGAG GACCAGAAGG GCACATCCA CTGAAACCT 1440

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70

TCCTTCTCCG ACGGGCTCAA GATGGACGCG GGCATCATCT GTGATGTGTG CACCTGGGAG 1500
 CTGCAAAAAG AGGTGCGGTC AGCTTCGCTG AGCTTCACAG GAGACTTCGT GTGCGGACAG 1560
 TGTGTGTGCA GCGAGGGCTG GAGTGGCCAG ACCTGCAACT GCTCCACCGG CTCTCTGAGT 1620
 GACATTCAAG CCTGCTTGGG GGAGGGGCGAG GACAAGCCGT GCTCGGGCCG TGGGAGTGC 1680
 CAGTGGGGGC ACTGTGTGTG CTACGGCGAA GGCGGCTACG AGGGTCAGTT CTGCGAGTAT 1740
 GACAACTTCC AGTGTCCCGG CACTTCGGGG TTCTCTGCA ATGACCGAGG AGCTGTCTCC 1800
 ATGGGCCAGT GTGTGTGTGA GCCTGGTTGG ACAGGCCCAA GCTGTGACTG TCCCTCAGC 1860
 AATGCCACCT GCATCGACAG CAATGGGGGC ATCTGTAATG GACGTGGCCA CTGTGAGTGT 1920
 GGGCGTGGC ACTGCCACCA GCAGTGGCTC TACACGGACA CCATCTGCGA GATCAACTAC 1980
 TGGCGATCC ACCCGGGCTT CTGCGAGGAC CTACGCTCCT GCGTGCAGTG CCAGGCGTGG 2040
 GGCACCGGGG AGAAGAAGGG GCGCACGTGT GAGGAATGCA ACTTCAAGGT CAAGATGGTG 2100
 GACGAGCTTA ACAGAGCCGA GGAGGTGGTG GTGCGCTGCT CCTTCCGGGA CGAGGATGAC 2160
 GACTGCACCT ACAGTACAC CATGGAAGGT GACGGGCCCC CTGGGCCCAA CAGCACTGTC 2220
 CTGGTGCACA AGAAGAAGGA CTGGCCCTCG GGCCTCTTCT GGTGGCTCAT CCCCTGCTC 2280
 TCCCTCTCTC TGCCGCTCCT GGCCTTGCTA CTGCTGCTAT GCTGGAAGTA CTGTGCTGTC 2340
 TGCAAGGCTT GCTTGGCAGT TCTCCCGTGC TGCAACCGAG GTCAATAGT GGGCTTTAAG 2400
 GAAGAGGCTT ACAGAGCTCT GAGGAACCTG ATGGCCTCTG ACCACTTGGA CACGCCCATG 2460
 CTGCGCAGCG GGAACCTTCA GGGCCGTGAC GTGGTCCGCT GGAAGGTAC CAACAACATG 2520
 CAGCGGCGTG GCTTTGGCCAC TCATGCGGCC AGCATCAACC CCACAGAGCT GGTGCGCTAC 2580
 GGGCTGTCTT TGCGCTTGGC CCGCTTTTGC ACCGAGAACC TGCTBAAGCC TGACACTCGG 2640
 GAGTGGCGCC AGCTGGCCCA GGAGGTGGAG GAGAACCCTG ACCAGGTCTA CAGGCAGATC 2700
 TCCGCTGTAT CAGAGCTTCA GAGACCCAG TTCCGGCAGC AGCCCAATGC CGGGAAGAAG 2760
 CAAGACCACA CCATTGTGGA CACAGTGTG ATGGCGCCCC GCTCGGCCAA GCGGGCCCTG 2820
 CTGAAGCTTA CAGAGAAGCA GGTGGAACAG AGGGCCCTTC ACBACCTCAA GGTGGCCCTC 2880
 GGTACTTACA CCTACTTCC AGACAGGAC GCGCGGGGCA TGGTGAAGT CCAGGAGGGC 2940
 GTGAGCTTGG TGAGAGTACG GTTGCCTCTC TTTATCCGGC CTGAGGATGA CGACGAGAG 3000
 CAGCTGTCTG TGGAGGCGAT CGAGCTGCCC GCAGGCACTG CCACCTCTCG CCGCGGCGTG 3060
 GTAAACATCA CCATCATCAA GGAGCAAGCC AGAGAGCTGG TGTCTTTGA GCAGCCTGAG 3120
 TTCTCGGTCA GCGCGGGGGA CCAGGTGGCC CGCATCCCTG TCATCCGGCG TGTCTGGGAC 3180
 GGGGGAAGT CCGAGGTCTC CTACCGACA CAGGATGGCA CCGCGCAGGG CAACCGGGAC 3240
 TACATCCCCG TGAGAGGTGA GCTGTGTTTC CAGCCTGGGG AGGCTCTGGA AGACGTGACG 3300
 GTGAAGCTTC TGAGAGCTGA AGAAGTTGAC TCCCTCTGTC GGGGCGGCCA GGTCCGCGTG 3360
 TTCCAGCTCC AGCTCAGCAA CCTTAAGTTT GGGGCCACCC TGGGCCAGCC CCACTCCACC 3420
 ACCATCATCA TCAGGAGCCC AGATGAAGTG GACCGGAGCT TCACGAGTCA GATGTTGTCA 3480
 TCACAGCCAC CCGCTCACGG CGACCTGGGC GCGCGGCGCA ACCCCATGCT TAAGGCGGCT 3540
 GGGTCCAGGA AGATCCATTT CAACTGGCTG CCCCCTCTCT GCAGGCAAT GGGGTACAGG 3600
 GTAAAGTAGT GGATTCAGGG TGACTCCGAA TCCGAHCC ACCTGCTCGA CAGCAAGGTG 3660
 CCGTCAGTGG AGCTCACCAA CCGTACCCG TATTGCGACT ATGAGATGAA GGTGTGCGCC 3720
 TACGGGGCTC AGGGCGAGGG ACCCTACAGC TCCCTGGTGT CCTGCCGAC CCACAGGAA 3780
 GTGGCGGCG AGCCAGGGCG TCTGGCCTTC AATGTGTGCT CCTCCACGGT GACCCAGCTG 3840
 AGCTGGGCTG AGCCGGCTGA GAOCACCGGT GAGATCAGAG CCTACGAGGT CTGCTATGGC 3900
 CTGTGTCAGG ATGACCAACC ACCTATTGGG CCAATGAAGA AAGTGTGTTT TGCAACCCCT 3960
 AAGAACCAGA TGCTGCTTAT TGAGAACCTT CCGGAGTCCC AGCCCTACCG CTACACGGTG 4020
 AAGGGCGGCA ACCGGGCGCG CTGGGGGCTT GAGCGGAGG CCATCATCAA CCTGGCCACC 4080
 CAGCCCAAGA GGGCCATGTC CATCCCATC ATCCCTGACA TCCCTATGCT GGACGCCAG 4140
 AGCGGGGAG ACTACGACAG CTTCCTTATG TACAGGATG ACCTTCTAAG CTCTCCATCG 4200
 GGCAGCCAGA GGGCTCAGGT CTCGATGAC ACTGAGCACC TGGTGAATGG CCGGATGGAG 4260
 TTTGCCCTTC CCGGCGACAC CACTCCCTG CACAGGATGA CCACGACAG TGCTGTGCTG 4320
 TATGGCACCC ACCTGAGCCC ACAGGTGCCC CACCGGCTGC TAAGCACATC CTCACCCCTC 4380
 ACACGGGACT ACAACTTACT GACCCGCTCA GACACTTAC ACTGACACAC ACTGCCGAGG 4440
 GACTACTTCA CCTTCACTTC CCGTCTCTCC CAGGACTCTC GCTTCACTGC TGGTGTGCCC 4500
 GACACGCCCA CCGGCTGCTG GTTCTCTGCT CTGGGGCCCA CATCTCTCAG AGTGAAGCTG 4560
 CAGAGGCGGC GGTGCGAGCG GCGCTGCGAG GGTACAGTG TGGAGTACCA GCTGCTGAAC 4620
 GGGGCTGAGC TGATCGGCTT CAACATCCCC AACCTGCCCC AGACCTCGGT GGTGGTGGAA 4680
 GACCTCTGTC CCAACCACTC CTAGGTGTTT CCGGTGCGGG CCCAGAGCCA GGAAGGCTGG 4740
 GGGCGAGAGC GTGAGGGTGT CATCACATT GATTOCCAGG TGACCCGCA GAGCCACTG 4800
 TGTCTCTGCT CAGGCTCCGC CTTCACCTTG AGCACTTCCA GTGCCCCAGG CCGGCTGTTG 4860
 TTCACTGCCC TGAGCCCGAG CTGCTGTCAG CTGAGCTGGG AGCGGCCAGG GAGGCCAAT 4920
 GGGGATATCG TGCGCTACCT GGTGACCTGT GAGATGGCCC AAGGAGGAGG GCCAGCCACC 4980
 GCATTCCGGG TGGATGGAGA CAGCCCGGAG AGCCGGCTGA CGGTGCGGG CCTCAGGAG 5040
 AAGTGGCCCT ACAAGTCAA GTTGCGAGCC AGGACCACTG AGGGCTTGG GCGAGAGCGC 5100
 GAGGGCATCA TCACCATAGA GTCCGAGGAT GGAAGGACCT TCCCGCAGCT GGGCAGCCGT 5160
 GCGGGGCTCT TCACGACCCC GCTGCAAGC GAGTACAGCA GCATCACAC CACCCACACC 5220
 AGGCCACCG AGCCCTTCTT AGTGGATGGG CCGACCTGG GGGCCAGCA CCTGGAGGCA 5280
 GGGGCTCTCC TCACCGGCA TGTGACCCAG GAGTTGTGA GCGGACACT GACCAACAGC 5340
 GGAACCTTCA GCACCCACAT GGACCAACAG TTCTTCCAA CTGACCCGA CCTGCCCCA 5400
 CCGCGGCTAT GTCCCACTAG GCGTCCCTCC GACTCCTCTC CCGGAGCCTC CTCAGCTACT 5460
 CCATCCTTGC ACCCTGSGGG GCGGAGCCCA CCGGATGCA CAGAGCAGGG GCTAGGTGTC 5520
 TCTGGGAGG CATGAAGGGG GCAAGGTCCG TCTCTGTGG GCGCAACCT ATTGTAAAC 5580
 AAAGAGCTGG GAGCAGCACA AGGACCCAGC CTTTGTCTG CACTTAATAA ATGGTTTGGC 5640
 TACTG

Seq ID NO: 360 Protein sequence
 Protein Accession #: NP_000204.1

75
 80

1 11 21 31 41 51
 MAGPRPSFWA RLILAALISV SLSGTLANRC NKAPVKSCTE CVRVKDCAY CTDEMFRDRR 60
 CNTQAEILAA GQRESIVVM ESSFQITEET QIDTLRRSQ MSPQGLRVRL RPEERHFEL 120
 EFFEPLSEPV DLYTLMDFSN SMSDDLDMLEK KMGQNLARVL SGLTSDYITG FGFVDKVSFV 180
 PQTMRPEKL KEFEPNSDFP FSEKNVISLT EDVDEFNKL QGKRISGNLD APEGGFDALE 240
 QTAQCTRIDG WEPSTHLLV PSTESAFHYE ADGANVLAGE MSRNDERCHL DTTGYTYQYR 300
 TDYPSVPTIL VRLAKENII PIFAVTNSY SYEKLHTYF PVSLGLVQS DSSNIVKLE 360
 EAFNRIRSNL DIRALDSFRG LRTEVTSKMF QKTRTGSFRI RRGSVGIYQV QLRALHVDG 420

THVQQLPEDQ KGNHLKPSF SDGLKMDAGI ICDVCTCELO KEVRSARCSF NGDFVCGQCV 480
 CSEGWSSGTC NCSTGSLSDI QPCLRGEGDK PCSGRGECQC GECVCYGBGR YEGQFCBYDN 540
 PQCPRTSGFL CNDRGRCSMG QCVCEPGWTG PSCDCPLSNA TCIDSNNGTC NRGHCECGR 600
 CHCQOSLYT DTICEINYSI IHPGLCEDLR SCVOCQAWGT GEKKGRTCEE CNFKVKMDE 660
 LKRAEEVVVR CSFRDEDDDC TYSYIMEGDS ARGPNSTVLV RKKKDCPPGS FWWLIPLLLL 720
 LLPLLALLLL LCMKYCACCK ACLALLPCCN RGHMVGFKED HYMLRENIMA SDHLDTPMLR 780
 SGNLNGRDVV RNKVTNNMQR PGPATHAASI NPTELVPYGL SLRLARLCTE NLKLPDTRC 840
 AQLRQVEEEN LNEVYRQISG VHLQQTQKFR QQPNAGKKQD HTIVDTVLMA PRSAKPALLK 900
 LTERQVEQRA FHDLKVAPGY YTLTADQDAR GMVEFQEGVE LVDVRVPLEFI RPEDDDEKQL 960
 LVEALDVVAG TATLGRRLVN ITILKEQARD VVSFPOPEFS VSRGDQVARI FVIRRLDGG 1020
 KSQVSYRTQD GTAQGNRDYI FVEGELLFOP GEAWKELQVK LLELQEVDEL LRGRQVRRFH 1080
 VQLSNPKFGR HLGQPHSTTI IIRDPDELDR SPTSOMLSQ PPPHGDIGAP QNPMKAAGS 1140
 RKIHFNNLFP SGKPMGYRVK YWIQGDESE AHLDSKVP SVELNLYPYC DYEMKVCAYG 1200
 AQSGEPYSSL VSCRTHQEVF SEPRIAFNV VSETVTQLSN AEPATNGEI TAYEVCYGLV 1260
 NDDNRPIGPM KKVLDVNPKN RMLLIENLRE SQPYRYTVKA RAGAGWGRER EAIINLATOP 1320
 KRPMSIPIIP DIFIVDAQSG EDYDSPLMYS DDVLRSPSGS QRPVSDDTE HLVNGRMDFA 1380
 FPGSTNSLGR MTTSAAYG THLSPHVPHR VLSTSTLTR DYNLSRBEH SHSTTLPRDY 1440
 SLTTSVSSD SRLTAVPBDT PTRLVPSALG PTLRLVSWQE PRCEPLOGY SVEYQLNGG 1500
 ELHRLNIPNP AQTGVVVEDL LPNHSYVFRV RAQSQBGWR EREGVITIEG QVHPQSLCP 1560
 LPGAFTLST PSAPGFLVFT ALSPDSLQLS WERPRRPNGD IVGYLVTCM AQGGGPATAF 1620
 RVDGDSPEER LTVPLGSENV FYKFKVQART TEGFGPEREG IITIESQDGG PFPQLGSRAG 1680
 LFQHPQLSEY SSITTTHTSA TEPFLVDGPT LGAQHLEAGG SLTRHVTOEF VSRITLTSGT 1740
 LSTHMDQQFF QT

Seq ID NO: 361 DNA sequence
 Nucleic Acid Accession #: NM_013332.1
 Coding sequence: 1..63

1 11 21 31 41 51
 GCACGAGGGC GCTTTGTCT CCGGTGAGTT TTGTGGCGGG AAGCTTCTGC GCTGGTGCTT 60
 AGTAACCGAC TTTCCTCCGG ACTCCTGCAC GACCTGCTCC TACAGCCGGC GATCCACTCC 120
 CGCGTGTTC CCCGAGGTT CCAGAGGCTC TTCAGAAGGA GAAGGCAGCT CTGTTTCTCT 180
 GCAGAGGAGT AGGTCCTTT CAGCATGAA GCATGTGTG AACCTCTACC TGTTAGGTGT 240
 GGTACTGACC CTACTCTCCA TCTTGGTAG AGTGATGAG TCCTTAGAAG GCTTACTAGA 300
 GAGCCCATCG CTGGGACCT CCTGGACAC CAGAAGCCAA CTAGCCAACA CAGAGCCAC 360
 CAAGGGCCTT CCAGACCATC CATCCAGAAG CATGTGATA GACCTCTTC CATACTGGCC 420
 ATATTTGGGA ACACTGACCT AGACATGTCC AGATGGGAGT CCAATTCTTA GCAGACAAGC 480
 TGAGCAACCT TGTAAACAGA GAACATTAC TAGGCTTGA AGAACCTGTC TAACTGGATG 540
 CTCATTGCTT GGCACAGGCC TGTTAGGCC GGTTCGGTG GCTCATGCCCT GTAATCCTAG 600
 CACTTTGGGA GGTCTGAGGT GGTGGATCAC CTGAGGTGAG GAGTTGAGA CCAGCCTCCG 660
 CACATGCGCG AAACCCCATC TCTACTAAAA ATACAAAGT TAGCTGGGTG TGGTGGCAGA 720
 GGCCTGTAAT CCGAGTTCTT TGGGAGGCTG AGGCGGGAGA ATTGCTTGAA CCGCGGGACG 780
 GAGGTTCGAG TGAACCGAGA TGCCACTGCT GTACCCAGCC TGGGCCACAG TGCAAGACTC 840
 CATCTCAAAA AAAAAAGAA AAGAAAAAGC CTGTTTAATG CACAGGTGTG AGTGGATTGC 900
 TTATGGCTAT GAGATAGGTT GATCTCGCCC TTACCCCGGG GTCTGGTGTG TGCTGTGCTT 960
 TCCTCAGCAG TATGGCTCTG ACATCTCTTA GATGTCCCAA CTTGAGCTGT TGGGAGATGG 1020
 TGATATTTTC AACCTTACTT CCTAAACATC TGCTGGGGT TCCTTTAGTC TTGAATGTCT 1080
 TATGCTCAAT TATTTGGGTG TGAGCCTCTC TTCCACAAGA GCTCCTCAT GTTTGGATAG 1140
 CAGTTGAGAA GGTGTGTGGG GTGGGCTGTT GGGAGTGAGG ATGGAGTGT CAGTGCCCAT 1200
 TTCTCATTTT ACATTTTAAA GTCGTTCCTC CAACATAGTG TGTATTGGTC TGRAGGGGGT 1260
 GGTGGGATGC CAAGCCTGTC TCAAGTTATG GACATTGTGG CCACCATGTG GCTTAAATGA 1320
 TTTTCTCTAA CTAATAAGT GGAATATATA TTTCAAAAA AAAAAAAAAA AA

Seq ID NO: 362 Protein sequence
 Protein Accession #: NP_037464.1

1 11 21 31 41 51
 MGHVLMNLYLL GVVLTLISIF VRVMSLEGL LSPSPGTSW TTRSQLANTE PTKSLPDHPS 60
 RSM

Seq ID NO: 363 DNA sequence
 Nucleic Acid Accession #: NM_023915.1
 Coding sequence: 250..1326

1 11 21 31 41 51
 GGCACGAGGG TTTCGTTTTC ATGCTTTACC AGAAAAATCCA CTTCCTGCCC GACCTTAGTT 60
 TCRAAGCTTA TTCTTAATTA GAGACAAGAA ACCTGTTTCA ACTTGAAGAC ACCGTATGAG 120
 GTGAATGGAC AGCCAGGCAC CACAATGAAA GAATCAAAAC CAGGAATAC CTATGCTGAA 180
 CCACGCGCTC AATCGTCCCC AAGTGTTTCC TGACAGCAT CTTTGTCTAC AGTGCAATCAC 240
 AACTGAAGAA TGGGGTTCAA CTGACGCTT GCAAAATTAC CAAATAACGA GCTGCACGAC 300
 CAAGAGATC ACAATTCAGG CAACAGGAGC AGCGGGCCAG GAAAGAACAC CACCTTAC 360
 AATGAATTTG ACNCAATTGT CTGCGCGTG CTTTATCTCA TTATATTTGT GGCAGCATC 420
 TTGCTGAATG GTTTAGCAGT GTGGATCTTC TTCCACATTA GGAATAAAC CAGCTTCATA 480
 TTCTATCTCA AAAACATAGT GGTTCAGAG CTCAATGGA CGCTGACATT TCATTTTGA 540
 ATAGTCCATG TTGACAGAIT TGGACCTTGG TACTTCAAGT TTATCTCTG CAGATACACT 600
 TCAGTTTTGT TTTATCAAA CATGTATACT TCCATCGTGT TCCTTGGGCT GATAAGCATT 660
 GATCGCTATC TGAAGTGGT CAAGCCATT GGGGACTCTC GGATGTACAG CATAACCTTC 720
 ACCGAGGTTT TATCTGTTTG TGTTTGGGTG ATCATGGCTG TTTTGTCTTT GCCAAACATC 780
 ATCTGACAA ATGCTCAGCC AACAGGGAC AATATCCATG ACTGCTCAAA ACTTAAAGT 840
 CTTTGGGGG TCAATGGCA TACGCGAGTC ACCTATGTGA ACAGCTGCTT GTTGTGBC 900
 GTGCTGTGTA TTCTGATCGG ATGTACATA GCCATATCCA GGTACATCCA CAAATCCAGC 960

AGGCAATTCA TAAGTCAGTC AAGCCGAAAG CGAAAACATA ACCAGAGCAT CAGGGTTGTT 1020
 GTGGCTGTGT TTTTACCTG CTTTCTACCA TATCACTTGT GCAGAAATCC TTTTACTTTT 1080
 AGTCACCTAG ACAGGCTTTT AGATGAATCT GCACAAAAA TCCTATATTA CTGCAAGAA 1140
 ATACACTTTT TCCTGTCTGC GTGTATGTT TCCCTGGATC CAATAATTTA CTTTTTCATG 1200
 TGTAGGTCAT TTCAAGAGG GCTGTTCAA AAATCAAATA TCAGAACCA GAGTGAAAGC 1260
 ATCAGATCAC TGCAAGTGT GAGAAGATCG GAAGTTCGCA TATATTATGA TTACACTGAT 1320
 GTGTAGGCTT TTTATTGTTT GTTGGAAATCG ATATGTACAA AGTGTAATAA AATGTTTCTT 1380
 TTCATTATCC TTAATAAAAA AA

Seq ID NO: 364 Protein sequence
 Protein Accession #: NP_076404

1 11 21 31 41 51
 | | | | |
 MGFNLTLAGL PNNELHQQES HNSGNREDGP GKNTTLHNEF DTIVLPVLYL IIFVASILIN 60
 GLAVWIFPHI RNKTSFIPLY KNIVVADLIM TLTFPFRIVH DAGFGPWYFK FILCRYTSVL 120
 FYANMTYSIV FLGLISIDRY LKVVKPFQDS RMYSTFTKV LSVCVWVIMA VLSLPNIILT 180
 NQOPTEDNII DCCKLKSLPLG VKNHTAVTYV NSCLFVAVLV ILIGCYLAIS RYIHKSSRQF 240
 ISQSSSRKHK NQSRVTVAVV FFTCFLPYHL CRIPFTFSLH DRLLDESAQK ILYYCKEITL 300
 FLSACSVCLD PIIFYFMCRS FSRRLFKKSN IRTRESIRS LQSVRRSEVR IYYDYTDV

Seq ID NO: 365 DNA sequence
 Nucleic Acid Accession #: NM_005365.1
 Coding sequence: 1..948

1 11 21 31 41 51
 | | | | |
 ATGTCCTCTCG AGCAGAGGAG TCCGCACTGC AAGCCTGATG AAGACCTTGA AGCCCAAGGA 60
 GAGGACTTGG GCCTGATGGG TGCACAGGAA CCCACAGGCG AGGAGGAGGA GACTACCTCC 120
 TCCCTCTGCA GCAAGGAGGA GGAGGTGTCT GCTGCTGGGT CATCAAGTCC TCCCAGAGT 180
 CCTCAGGGAG GCGCTTCCTC CTCCATTTCC GTCTACTACA CTTTATGGAG CCAATTGGAT 240
 GAGGGCTCCA GCAGTCAAGA AGAGGAAGAG CCAAGCTCCT CGGTGACCCC AGCTCAGCTG 300
 GAGTTCATBT TCCAAGAGCG ACTGAATTTG AAGGTGGCTG AGTTGGTTCA TTTCTGCTC 360
 CACAAATATC GAGTCRAGGA GCCGGTCACA AAGGCAGAAA TGCTGGAGAG CGTCATCAAA 420
 AATTACAAGC GCTACTTTCC TGTGATCTTC GGCAGAGCCT CCGAGTTTCA GCAGGTGATC 480
 TTTGGCACTG ATGTGAAGGA GGTGGACCCC GCGGCCCCCT CCTACATCCT TGTCACTGCT 540
 CTGCGCTCTT CGTGGCATAG CATGCTGGGT GATGGTCATA GCATGCCCAA GCGCGCCCTC 600
 CTGATCATTT TCTTGGTGTG GATCCTAACC AAGACAACT GCGCCCTTGA AGAGGTTATC 660
 TGGGAAGCOT TGAGTGTGAT GCGGGTGTAT GTTGGGAAGG AGCACAATGT CTACGGGAG 720
 CCGAGGAGCG TGCTACCCCA AGATTGGGTG CAGGAAACT ACCTGGAGTA CCGGCAGGTG 780
 CCGGCACTG ATCTTCGCGA CTACGAGTTC CTGTGGGCTT CCAAGGCCCA CGCTGAATC 840
 AGCTATGAGA AGGTCATAAA TTATTGGTTC ATGCTCAATG CAAGAGAGCC CATCTGCTAC 900
 CCTCCCTTT ATGAAGAGGT TTTGGGAGAG GAGCAAGAGG GAGTCTGA

Seq ID NO: 366 Protein sequence
 Protein Accession #: NP_005356.1

1 11 21 31 41 51
 | | | | |
 MSLSQRSPHC KPDRLDAQE EDLGLMGAQE PTGEEETTS SSSKEKEVS AAGSSSPFQS 60
 PQGASSSIS VYTTLSQFD EGSSSQEERE PSSVDPAQL ERMFORALKL KVABLVRHLL 120
 HKYRVKEPVT KAEMLSVVIK NYKRYFPVIF GKASEFMQVI PGTDVKEVDP AGHSYILVTA 180
 LGLSCDSMLG DGHSMPLKAL LIIIVLGVIIT KDNCAPEEVI WEALSVMGVY VGKEMFVGR 240
 FRKLLTDQWV QENYLEYRQV PGSDPAHYEF LMGSKAHAEI SYEKVINYLV MLNAREPICY 300
 PSLYEELVGE EQEGV

Seq ID NO: 367 DNA sequence
 Nucleic Acid Accession #: NM_014400
 Coding sequence: 86..1126

1 11 21 31 41 51
 | | | | |
 GGTACTCAT CCTGGGCTCA GGTAAAGAGG CCCAGCTCG GAGGCGGCAC ACCCAGGGGG 60
 GACGCCAAGG GAGCAGGACG GAGCCATGGA CCCCAGGAGG AAGCAGGTG CCCAGGCCAT 120
 GATCTGGACT CGAGGCTGGC TGCTGCTGCT GCTGCTTCGC GAGGAGGCGC AGGCCCTGGA 180
 GTGCTACAGC TGCGTGCAGA AAGCAGATGA CGGATGCTCC CGGAACAAGA TGAAGACAGT 240
 GAGGTGCGCG CCGGCGGTGG AGCTCTGCAC CGAGGCGGTG GGGGCGGTGG AGACCATCCA 300
 CGGACAATTC TGCTGGCAG TGCGGGGTTG CGGTTGCGGA CTCGCCGCGA AGAATGACCG 360
 CGGCTGGAT CTTCACGGGC TTCTGGCGTT CATCCAGCTG CAGCAATGCG CTCAGGATCG 420
 CTGCAACGCC AAGCTCAACC TCACTGCGCG GCGCTCGAC CCGGCAAGTA ATGAGAGTGC 480
 ATACCGCGCC AACGGCGTGG AGTGTACAG CTGTGTGGC CTGAGCCGCG AGGCGTGCCA 540
 GGGTACATCG CCGCGCGTGC TGAAGTGTCT CAACGCCAGC GATCATGTCT ACAAGGGCTG 600
 CTTCGACGCG AACGTCACTT TGAAGGAGCG TAATGTGACT GTGTCTTTCG CTGTCCGGGG 660
 CTGTGTCCAG ATGCAATTCT GCATCGGGA TGGAGTAACA GGCACAGGGT TCAGGCTCAG 720
 TGGCTCTCTG TGCCAGGGGT CCGCTGTAA CTCTGACCTC CGCAACRAGA CCTACTTCTC 780
 CCTCGAATC CTCTGTGTTG TCGGCTGCG CCGTCCAGAG CCCAGACTG TGGCCTCAAC 840
 CACATCTGTC ACCACTTTCT CCTCGGCCCG AGTGAGACCC ACATCCACCA CCAAAACCAT 900
 GCTAGCGCCA ACCAGTCAGA CTCGAGAGCA GGGAGTAGAA CACGAGGCGT CCGGGATGTA 960
 GAGAGCCAGG TGTGATGAG GCGCGCTGG CACACAGGAC CGCAGCAATT CAGGGCAGTA 1020
 TCCGCAAAA GGGGGGGCCC AGCAGGCCCA TAATAAAGGC TGTGTGGCTC CCACAGCTGG 1080
 ATTGGCAGC CTCTGTGTTG CGTGGCTGCT TGGTGTCTTA CTGTGAGCTT CTCACCTGCG 1140
 AAATTTCCCT CTCACCTACT TCTCTGGGCC TGGGTACCCC TCTTCTCATC ACTTCTCTGT 1200
 CCCACCACTG GACTGGGCTG GCGGAGCCCC TGTTTTCCCA ACATTCGCCA GTATCCCCAG 1260
 CTTCTGCTGC GCTGTTTGC GCGTTTGGGA AATAAATAC CGTTGTATAT ATTCTGGCAG 1320

GGGTGTTCTA GCTTTTGTAG GACAGCTCCT GTATCCTTCT CATCCTTGTC TCTCCGCTTG 1380
 TCCTCTTGTG ATGTTAGGAC AGAGTGAGAG AAGTCAGCTG TCACGGGGAA GGTGAGAGAG 1440
 AGGATGCTAA GCTTCCCTACT CACTTTCTCC TAGCCAGCCT GGACTTTGGA GCGTGGGGTG 1500
 5 GGTGGGACAA TGGCTCCCCA CTCTAAGCAC TGCCCTCCCT ACTCCCGCA TCTTTGGGGA 1560
 ATCGGTTCCT CATATGTCTT CCTTACTAGA CTGTGAGCTC CTCGAGGGCA GGGACCGTGC 1620
 CTTATGCTTG TGTGTGATCA GTTTCTGGCA CATAAATGCC TCATAAAGA TTTAATTACT 1680
 TTGTATAGTG AAAAAAAA

Seq ID NO: 368 Protein sequence
 Protein Accession #: NP_055215

1 11 21 31 41 51
 15 MDPARKAGAG AMIWTAGNLL LLLLRGGAQA LECYSCVQKA DDGCSFNKMK TVKCAPGVDD 60
 CTEAVGAVET IHGQFSLAVX GCGSGLPKKN DRGLDLHGLL AFIQLQQCAQ DRGMAKLNLT 120
 SRALDPAGNE SAYPPNGVEC YSCVGLSREA CQGTSPFVVS CYNASDHVYK GCPDGNVTLT 180
 AANVTVSLPV RGCVDDEFT RDGVTGPQFT LSGSCQGSR CNSDLRNKTY FSPRIPLVLR 240
 LPPEPTTFA STTSVTTSTF AFVPTSTTK PMPAPTSTQF RQGVHEASR DEEPRLTGGA 300
 20 AGHQDRSNG QYPAKGGPQQ PHNKGCVAPT AGLAALLLAV AAGVLL

Seq ID NO: 369 DNA sequence
 Nucleic Acid Accession #: NM_005329.1
 Coding sequence: 1..1662

25 1 11 21 31 41 51
 30 ATGCCGCTGC AGCTGACGAC AGCCCTGCGT GTGCTGGGCA CCAGCCTGTT TGCCCTGGCA 60
 GTGCTGGGCT GCATCCTGGC AGCCTATGTG ACGGCTACCC AGTTTCATCA CACGGAAAAG 120
 CACTACCTGT CTTTCGGCCT GTACGGCGCC ATCCTGGGCC TGCACTTGCT CATTCAGAGC 180
 CTTTTCGCTT TCCTGGAGCA CGGGCGCATG CGAGCTGCTG GCCAGGCCCT GAAGCTGCCC 240
 TCCTCCGCGC GGGGCTCGGT GGCCTCTGTC ATTGCCGCAT ACCAGGAGGA CCCTGACTAC 300
 TTGGCAAGT GCTTGGCTTC GCGCCAGCAG ATCTCCTTCC CTGACCTCAA GGTGGTCAIG 360
 GTGGTGGATG GCAACCGCCA GAGGAGCGCC TACATGCTGG ACATCTTCCA CGAGGTGCTG 420
 GCGGCGACCG AGCAGGCCCG CTTCCTTGTG TGGCGCAGCA ACTTCCATGA GGCAGCGCAG 480
 35 GGTGAGACGG AGGCCAGCCT GCAGGAGGGC ATGAGACCGT TCGCGGATGT GGTGCGGGCC 540
 AGCACTTCTT CGTGCTCAT GCAGAGGTGG GAGGCGAAGC GCGAGGTGAT GTACACGGCC 600
 TTCAAGGCCG TCAGCGATTC GGTGGACTAC ATCCAGGTGT GCGACTCTGA CACTGTGCTG 660
 GATCCAGCCT GCACCATCGA GATGCTTCGA GTCTGGAGG AGGATCCCCA AGTAGGGGGA 720
 40 GTGCGGGGAG ATGTCCAGAT CCTCAACAAG TACCACTCAT GGATTTCCCT CCTGAGCAGC 780
 GTGCGGTACT GGATGGCCTT CAACGTGGAG CGGGCCTGCC AGTCTTACTT TGGCTGTGTG 840
 CAGTGTATTA TTGGGCCCCT GGGCATGTAC CGCAACAGCC TCCTCCAGCA GTTCTCTGGG 900
 GACTGTATCC ATCAAGAATT CTAAGCAGC AAGTGCAGCT TCGGGGATGA CCGGCACTTC 960
 ACCAAGCCAG TCTTGAGCCT TGGCTACCGA ACTAAGTATA CCGCGCGCTC CAAGTGGCTC 1020
 45 ACAGAGACCC CCACTAAGTA CTTCCGGTGG CTCAACAGC AAACCGCTG GAGCAAGTCT 1080
 TACTTCGGGG AGTGGCTCTA CAACTCTCTG TGGTTCCATA AGCAACCACCT CTGGATGACC 1140
 TACAGATCAG TGGTCAAGCG TTCTTCCCTC TTCTTCTTCA TTGCCACGGT TATACAGCTT 1200
 TTCTACCGGG GCGGCACTTC GAACATTCTC CTCTTCTGCG TGACGGTGCA GCTGGTGGGC 1260
 ATTATCAGGG CCACCTACGC CTGCTTCCCT CCGGGCAATG CAGAGATGAT CTTTCTGTCC 1320
 50 CTCTACTCCC TCCCTATATAT GTCCAGCCTT CTGCGCGCCA AGATCTTTCG CATTGCTACC 1380
 ATCAACAAAT CTGGCTGGGG CACCTCTGGC CAAAAACCA TTGTGGTGAA CTTCTATGGC 1440
 CTCATTCTCG TGTCCATCTG GGTGGCAGTT CTCTGGAGG GGCCTGGCTA CACAGCTTAT 1500
 TGCCAGGACC TGTTCAGTGA GACAGAGCTA GCGCTTCTTG TCTCTGGGCG TATACTGTAT 1560
 55 GGCTCTACT GGTGGGCCCT CCTCAGCTA TATCTGGCCA TCATCGCCCG GCGATGTGGG 1620
 AAGAAGCCGG AGCAGTACAG CTTGGCTTTT GCTGAGGTGT GA

Seq ID NO: 370 Protein sequence
 Protein Accession #: NP_005320.1

60 1 11 21 31 41 51
 MPVQLTALR VVGTSLFALA VLGGTLAAYV TGYQFIETEK HYLSPGLYGA ILHLRLLIQS 60
 LFAFLHRRM RRAGQALKLP SFRGGSVALC IAAYQEDPDY LRKCLRSQR ISFPDLKVVV 120
 VVDGNRQEDA YMLDIFHEVL GGTGAGGFV WRSNFHEAGE GETEASLQEG MDRVRDVVRA 180
 65 STPSCIMQXW GCKREVMYTA FKALGDSVDY IQVCDSDTVL DPACTIEMLR VLEBDPQVGG 240
 VGGDVQILNK YDSWISFLSS VRYWMAFNVE RACQSYFGCV QCISGPIGMY RNSLLQQFLR 300
 DWYHQFLGS KCSFGDDRHL TNRVLSLGYR TRYTARSKCL TETPFKYLWR LNQQTRWSKS 360
 YFRWLYNSL WFKHHLWMT YESVVTGFFP FFLIATVIQL FYRGRWNL LFLLTQVLVG 420
 IIRATYACFL RGNAMIFMS LYSLLYMSBL LPAKIFALAT INKSGWGTSG RKTIVVMFIG 480
 70 LIPVSIWVAV LLEGAYTAY CQDLFSSTEL AFLVSGAILY GCYNVALIML YLAIXARROG 540
 KRPEQISLAF AEV

Seq ID NO: 371 DNA sequence
 Nucleic Acid Accession #: Eos sequence
 Coding sequence: 148-7095

75 1 11 21 31 41 51
 80 CACACATACG CACGACCGAT CTCACCTCGA TCTATACACT GGAGGATTAA AACAAACAAA 60
 CAAAAAAGAC APTTCCTTGG CTCCCTCCCT CTCTCCACTC TGAGAGCAG AGGAGCCGCA 120
 CGGCGAGGGG CCGCAGACCG TCTGGAAATG CGAATCCTAA AGCGTTTCTT CGCTTGCAAT 180
 CAGCTCCTCT GTGTTTGGCG CCGTGGATTG GCTAATGGAT ACTACAGACA ACAGAGAAAA 240
 CTTGTTGAAG AGATTGGCTG GTCCATATCA GAGCACTGA ATCAAAAAAA TTGGGGAAAG 300
 AAATATCCAA CATGTAATAG CCCAAACAAA TCTCCTATCA ATATTGATGA AGATCTTACA 360
 CAAGTAAATG TGAATCTTAA GAAACTTAAA TTTGAGGTT GGGATAAAC ATCATTGGA 420

	AACACATTCA	TTTATAACAC	TGGGAAAACA	GTGGAAATTA	ATCTCATTAA	TGACTACCGT	480
	GTCCAGCGAG	GAGTTTCABA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
	AAATGCAATA	TGTCTATCTGA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTCCTT	GTTTGGAGGA	AGCAGTCAAA	660
5	GGAAAAGGGA	AGTTAAAGAGC	TTTATCCATT	TGTGTTGAGG	TGGGACAGA	AGAAAATTG	720
	GATTTCAAAG	CGATTATTGA	TGGAGTCGAA	AGTGTAGTC	GTTTGGGAA	GCAGCTGCT	780
	TTAGATCCAT	TCATACTGTT	GAACCTTCG	CCAAACTCAA	CTGACAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCTGACCA	GACACAGTTG	ACTGGATTGT	TTTAAAGAT	900
10	ACAGTTAGCA	TCTCTGAAGG	CCAGTTGGCT	GTTTTTGTG	AAGTTCTTAC	AATGCAACAA	960
	TCTGTTATG	TCATGCTCAT	GGACTACTTA	CAAAACAATT	TTCCAGAGCA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTTC	CTCATACACT	GGAAAGGAAG	AGATTTCATG	AGCAGTTTGT	1080
	AGTTCCAGAC	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAGCCT	TCTGTGTACA	1140
	TGGGAAGAC	CTCAGTCTGT	TTATGATACC	ATGATTGAGA	AGTTTGCAGT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTGA	CAGATGGCTA	TCAAGACTTG	1260
15	GGTGTATTTC	TCAATAAATT	GCTAGCCCAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACATATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATTCCTGAAC	TTGATCTTTT	CCCTGAAATTA	ATTGGAACGT	AAGAAATAT	CAGAGGAGG	1440
	GAGAGGGGAA	AAGACATTGA	AGAAGGCGCT	ATGTGAAATC	CTGGTAGAGA	CAGTGTCTACA	1500
	AACCAAAATC	GGAAAAGGA	ACCCGAGATT	TCTACCACAA	CACACTACAA	TCCGATAGGG	1560
20	ACGAATACAC	ATGAAGCCAA	BACTAACCGA	TCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGAATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAGG	ATATTCTCTT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCACACTGCT	1740
	GAGGTACTTT	CAGCTCTCTT	AAATGATGGC	TCTAAACTCT	TTCTTAGATC	TCCCATATG	1800
	AACCTGTCTG	GAGAGTACCA	ATCTTAAAT	ACAGTTTCTA	TAACAGATA	TGAGGAGGAG	1860
25	AGTTTATTGA	CCAGTTTCAA	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAT	CATCTCTGAG	AACATATCCC	AAGGGTATAT	ATTTTCTCTC	1980
	GAAAACCCAG	AGACATAAAC	ATATGATGTC	CTTATACCGA	AATCTGCTAT	AAATGCTTCC	2040
	GAGATTTCAT	CTTCTATCAG	TTTCAAGGAA	TCACTAAGAG	ATCCTTCTAT	GGAGGGAAAT	2100
	GTGTGTTTCC	CTAGCTCTAC	AGACATAACA	GCACAGCCCG	ATGTTGGATC	AGGCAGAGAT	2160
30	AGCTTTCTCC	AGACTAATTA	CACGTAGATA	CGTGTGATG	AATCTGAGAA	GACAACCAAG	2220
	TCTTTTCTG	CAGGCCCATG	GATGTACAG	GGTCCCTCAG	TTACAGATCT	GGAAATGCCA	2280
	CATATCTCTA	CTTTTGCCCTA	CTTCCCAACT	GAGGTAAACAC	CTCATGCTTT	TACCCCATCC	2340
	TCCAGACAAC	AGGATTGGGT	CTTCAAGGTC	AACGTGGTAT	ACTCGCAGAC	AACCCACCCG	2400
	GTATACATAT	GTGAGACACC	TCTTCAACCT	TCTTCAAGTA	GTGAAGTCTT	TCCTCTAGTC	2460
35	ACCCCTTTGT	TGCTTGACAA	TCAGATCTTC	AACACTACCC	CTGCTGCTTC	AAGTAGTGAT	2520
	TCCGCCCTTG	ATGCTACGCC	TGTATTTCCC	AGTGTGATG	TGTCATTTGA	ATCCATCTCT	2580
	TCTTCCATAT	ATGCTGACCC	TTTGGCTTCA	TTTTCCTCTG	CTTCCCTCAG	TAGTGAATTG	2640
	TTTCCCATC	TGCTATACAT	TTCTCAAATC	CTTCCCAAG	TTACTTCTAG	TACCCAGAGT	2700
40	GATAAGTTGC	CTTCTCTATG	TTCTCTGCTA	GTGGCTGGGG	GTGATTGCTT	ATTAGAGCCC	2760
	AGCCCTTGCT	AGTATTTCTA	TGTGCTGTCC	ACTACTCATG	CTGCTTCAGA	GAOCTGGAA	2820
	TTTGGTAGTG	AACTCTGGTG	TCTTTATATA	ACGCTTATGT	TTTCTCAAGT	TGAACCAACC	2880
	AGCAGTATAT	CCATGATGCA	TGCAAGTTCT	TCAGGGCCCTG	AACCTTCTTA	TGCTTGTCT	2940
	GATAAATGAG	GCTTCCCAACA	CATCTTCACT	GTTCCTTACA	GTTCCTGCAAT	ACCTGTGCTAT	3000
	GATTCGTGG	GTGATCTCTA	TCAGGGTTCC	TTATTTAGCG	GCCCTAGCCA	TATACCAATA	3060
45	CCTAAGTCTT	CGTTAATAAC	CCCAACTGCA	TCATTACTGC	AGCCTACTCA	TGCCCTCTCT	3120
	GGTGATGGGG	AAAGGTCTGG	AGCCCTCTCT	GATAGTGAAT	TTCTTTTACC	TGACACAGAT	3180
	GGGCTGACAG	CCCTTAACAT	TTCTTCACTT	GTTCCTGATG	CTGAATTTAC	ATATACAAAC	3240
	TCTGTGTTTG	GTGATGATTA	TAAAGCGCTT	TCTAAAGTGT	AAATAATATA	TGGAATGAG	3300
	ACTGAACATG	CAATCTCTCT	TTTCAATGAG	ATGGTTTACC	CTTCTGAAAG	CACAGTCTAT	3360
50	CCCAACATGT	ATGATTAATGT	AAATAAGTGT	AATGCTCTT	TACAAGAAAC	CTCTGTTTCC	3420
	ATTTCTAGCA	CCAAGGCGAT	GTTTCCAGGG	TCCCTTGCTC	ATACCACCAC	TAAGGTTTTT	3480
	GATCATGAGA	TTAGTCAAGT	TCCAGAAAT	AACTTTTCAG	TTCAACCTAC	ACATAGTCTC	3540
	TCTCAAGCAT	CTGGTGACAC	TTCCGCTTAA	CCTGTGCTTA	GTGCAAACTC	AGAGCCAGCA	3600
	TCTCTGACCT	CTGCTTCTAG	TGAAGTGTTA	TCTCTTCAA	CTCAGCTCTT	ATTTTATGAG	3660
55	ACCTCAGCTT	CTTTTAGTAC	TGAAGTATG	CTCAACCTT	CCTTTCAGGC	TTCTGATGTT	3720
	GACACCTTGC	TTAAACTGT	TCTTCCAGCT	GTGCCAGTG	ATCCAAZATT	GTTTGAACCC	3780
	CCCAAGGTG	ATAAAATTAG	TTCTACAATG	TTGATCTTCA	TTGATCAAA	TTCTGCTTCA	3840
	AGTGAAGAAC	TGCTGCACTC	TACATCTGTA	CCAGTTTTCG	ATGTGTGCCC	TACTTCTCAT	3900
	ATGCTACTCTG	CTTCACTTCA	AGGTTTGAAC	ATTTCTTATG	CNAGTGAGAA	ATATGAACCA	3960
60	GTTTTGTTAA	AAAGTGAAG	TTCCACCAA	GTGGTAOCTT	CTTTGTACAG	TAATGATGAG	4020
	TTGTTCCAAA	CGGCCAATTT	GGAGATTAA	CAGGCCCTATC	CCCCAAAGG	AAGGCATGTA	4080
	TTTGCTACAC	CTGTTTTATC	AATTGATGAA	CCATTAAATA	CACATAATAA	TAAGCTTATA	4140
	CATTCCGATG	AAATTTTAA	CTCCACCAA	AGTTCTGTTA	CTGTTAAGGT	ATTTGCTGCT	4200
	ATTTCAACAG	TTGCTTCTGA	TACATTTGTA	TCTACTGATC	ATTTCTGTTCC	TATAGGAAT	4260
65	GGGCTAGTTG	CCATTACAGC	TGTTTCTCCC	CACAGAGATG	GTTCGTATAC	CTCAACAAAG	4320
	TTGCTGTCTC	CTTCTAAGGC	AACCTTCTGAG	CTGAGTCATA	GTGCCAAATC	TGATGCGGCT	4380
	TTAGTGGGTG	GTGGTGAAGA	TGGTGACACT	GATGATGATG	GTGATGATGA	TGATGATGAC	4440
	AGAGGTAGTG	ATGGCTTATC	CATTCTAAG	TGATGTCTAT	GCTCATOCTA	TAGGAATCA	4500
70	CAGGAAAAGG	TAAATGAATGA	TTCCAGACCC	CACGAAAACA	GTCTTATGGA	TCAGATAAT	4560
	CCATCTCTAT	ACTCACTATC	TGAGAAATCT	GAAGAGATA	ATAGAGTCAC	AAATGATATC	4620
	TCAGACAGTC	AAACTGGTAT	GGACAGAGT	CCTGGTAAAT	CACCATCAGC	AAATGGGCTA	4680
	TCCCAAAAGC	ACAATGATGG	AAAAGAGGAA	AATGACATTC	AGACGTGTAG	TGCTCTGCTT	4740
	CCTCTCAGCC	CTGAATCTTAA	AGCATGGGCA	GTTCGTACAA	GTGATGAAGA	AAATGGATCA	4800
	GGGCAAGGTA	TCTGAGTAG	CCCTAATGAG	AATGAGACTT	CCACAGATTT	CAGTTTGTGA	4860
75	GACACTAATG	AAAAGATGC	TGATGGGATC	CTGGCAGCAG	GTGACTCAGA	ATAACTTCCT	4920
	GGATTCCTAC	AGTCCCCAAC	ATCATCTGTT	ACTAGCGAGA	ACTCAGAGT	GTTCCACGTT	4980
	TCAGAGGCGAG	AGGCCAGTAA	TAGTAGCCAT	GAGTCTGTGA	TTGGTCTAGC	TGAGGGGTTG	5040
	GAATCCGAGA	AGAAAGGCGT	TATACCCCTT	GATGCTGTGT	CAGCCCTGAC	TTTTATCTGT	5100
	CTAGTGGTTC	TTTGGGTTAT	TCTCATCTAC	TGGAGGAAAT	GCTTCCAGAC	TGCACTCTTT	5160
	TACTTAGAGG	ACAGTACATC	CCCTAGAGTT	ATATCCACAC	CTCCACACCT	TATCTTCCA	5220
80	ATTTCAAGATG	ATGTCTGAGC	AATTCRAATA	AAGCACTTTC	CAGAGCATGT	TGCAGATTTA	5280
	CATGCAAGTA	GTGGGTTTAC	TGAAGAAATT	GAGACACTGA	AAGAGTTTAA	CCAGGAAGTG	5340
	CAGAGCTGTA	CTGTTGACTT	AGGTATTAACA	GCAGACAGCT	CCAACCACCC	AGACACAGAG	5400
	CACAAGAAATC	GATACATAAA	TATGTTTCC	TATGATCATA	GCAGGGTTAA	GCTAGCACAG	5460

5 CTTGCTGAAA AGGATGGCAA ACTGACTGAT TATATCAATG CCAATTATGT TGATGGCTAC 5520
 AACAGACCAA AAGCTTATAT TGCTGCCCAA GGCCCACTGA AATCCACAGC TGAAGATTTC 5580
 TGGAGAATGA TATGGGAACA TAATGTGAGG GTTATTGTCA TGATAACAAA CCTCGTGGAG 5640
 AAGGAGAGGA GAATATGTGA TCAGTACTGG CCTGCCGATG GGAGTGAGGA GTACGGGAAC 5700
 10 TTTCTGGTCA CTGAGAAGAG TGTGCAAGTG CTTGCCCTAT ATACTGTGAG GAATTTTACT 5760
 CTAAGAAACA CAAAATAAAA AAAGGGCTCC CAGAAAGGAA GACCCAGTGG ACGTGTGGTC 5820
 ACACAGTATC ACTACACGCA GTGCCCTGAC ATGGGAGTAC CAGAGTACTC CCTGCCAGTG 5880
 CTGACCTTTG TGAGAAAGGC AGCCTATGCC AAGCGCCATG CAGTGGGGCC TGTGTCTGTC 5940
 CACTGCAGTG CTGGAGTTGG AAGAACAGGC ACATATATTG TGCTAGACAG TATGTTGCAG 6000
 CAGATTCAAC ACGAAGGAAC TGTCAACATA TTTGGCTTCT TAAACACAT CCGTTCACAA 6060
 AGAATATTAT TGGTACAAAC TGAGGAGCAA TATGTCITCA TTCATGATAC ACTGGTTGAG 6120
 GGCATACCTA GTAAAGAAAC TGAGGTGCTG GACAGTCATA TTCATGCCCTA TGTTAATGCA 6180
 CTCTCATTC CTGGACCAGC AGGCAAAACA AAGCTAGAGA AACAAATCCA GCTCCTGAGC 6240
 CAGTCAATA TACAGCAGAG TGAATATTCT GCAGCCCTAA AGCAATGCAA CAGGGAAGAG 6300
 15 AATCGAACTT CTTCTATCAT CCCTGTGAAA AGATCAAGGG TTGGCATTTT ATCCCTGAGT 6360
 GGAGAGGGCA CAGACTACAT CAATGCCCTCC TATATCATGG GCTATTACCA GAGCAATGAA 6420
 TTCAATCATTA CCCAGCACCC TCTCCTTCAT ACCATCAAGG ATTTCTGGAG GATGATATGG 6480
 GACCAATAAT CCCAATGGT GGTATGATTT OCTGATGGCC AAAACATGGC AGAAGATGAA 6540
 TTTGTTTACT GGCCAAATGA AGATGAGCCT ATAAATTTGT AGAGCTTTAA GGTCACTTCT 6600
 20 AITGGCTGAG AACACAAATG TCTATCTAAT GAGGAAAAAC TTATAATTCA GGCATTTATC 6660
 TTAGAAGCTA CACAGAGTGA TTATGTACTT GAGGTGAGGC ACTTTCAGTG TCCTAAATGG 6720
 CCAATCCAG ATAGCCCCAT TAGTAAAACT TTTGAACCTA TAAGTGTTAT AAAAGRAGRA 6780
 GCTGCCAATA GGGATGGGCC TATGATTGTT CATGATGAGC ATGGAGGAGT GACGGCAGGA 6840
 ACTTCTGTTG TCTCGACAAC CCTTATGCAC CAACAGAAA AAGAAATTC CGTGGATGTT 6900
 25 TACCGGTAG CCAAGATGAT CAATCTGATG AGGCCAGGAG TCTTTGCTGA CATTGAGCAG 6960
 TATCAGTTTC TCTACAAAGT GATCCTCAGC CTGTGTGAGCA CAGGCGAGGA AGAGATCCA 7020
 TCCACCTCTC TGGACAGTAA TGGTGCAGCA TTGCTGATG GAAATATAGC TGAGAGCTTA 7080
 GAGTCTTTAG TTTAACACAG AAAGGGGGTG GGGGACTCAC ATCTGAGCAT TGTFTTCTC 7140
 TTCTAAATG TAGGCCAAGG AATCAGTCTA GTTCTGTTAT CTGTTGATT CCATCACCTT 7200
 30 GACAGTAAC TTCTATGAT AGGATTCTGC CGCCAAATTT ATATCATTAA CAATGTGTGC 7260
 CTTTTTCGAA GACTTGTAT TACTTATTA TGTGTGAAT AAAATGATTG AATTTTACAG 7320
 TATTCTAAG AATGGAATG TGTATTTT TCTGTATTG ATTTTACAG AAAATTTCAA 7380
 TTTATAGAGG TTAGGAATTC CAACTACAG AAAATGTTT TTTTATGAT CAAATTTTAA 7440
 GCTGATTTG TAGCAATAT CAGGTTTGCT AGAATATAA CTTTAAATC AGTAGCCCTG 7500
 35 AAATAAACA CTCTCCATA TGATATTCAA CATTTTACAA CTGCAGTATT CACCTAAAGT 7560
 AGAATAATC TGTACTTAT TGTAATACT GCCCTAGTGT CTCCATGGAC CAAATTTATA 7620
 TTTATAATTG TATTTTACTA TATTTTACTA CTGAGTCAAG TTTTCTAGTT CTGTGTATTT 7680
 GTTAGTTTAT ATGACGTAGT TCATTAGCTG GTCTTACTCT ACCAGTTTTC TGACATTGTA 7740
 TGTGTTACC TAGTCTATTA ACTTTGTTTC AGCATGTAAT TTTAACTTTT GTGGAATAA 7800
 40 GAAATACCTT CATTTTGAAG GAAGTTTFTA TGAGAATAAC ACCTTACCAA ACATGTGTTA 7860
 AATGTTTTTT ATCCAAGGAA TTGCAAAAT AAATATAAAT ATTGCCATTA AAAAAAATA 7920
 AAAAAAATA AAAAAAATA AAAA

Seq ID NO: 372 Protein sequence:

Protein Accession #: built from XP_031379

45
 1 11 21 31 41 51
 50 MRILKRFLAC IQLLCVCRID WANGYYRQQR KLVEEIGWSY TGALEQXNWG KKYPTCNSPK 60
 QSEINIDEDL TQVNVNKLKL KFQWDRKISL ENTFIENTGK TVEINLTNDY RVSGGVSEMV 120
 FKASKITPFW GKCNMSDGS EHSLEGQKFP LEMQIYCFDA DRPSSFEBAV KKGKLRALS 180
 ILFEVGTEN LDFKAIIDGV ESVSRFQKQA ALDPFILLNL LPNSTDKYI YNGSLTSPPC 240
 TDTVDNIVFK DTVESSESQL AVFCEVLTMQ QGGVYVLMNDY LQNNFRBQQY KPSRQVFSY 300
 TSKREIHEAV CSSEFENVQA DPENYTSLLV TWNERPRVVD TMIEKFAVLY QQLDGEDQTK 360
 55 HEPLTDGYQD LGAILNLLLP NMSVVLQIVA ICTNGLYGKY SDQLIVDMPT DNPRLDLFPE 420
 LIGTEIIEE ESEKGDIEEG AIVMPGRDSA TNQIRKBPQ IGSTTHYMRI GIKYNEAKTN 480
 RSPTRGSEFS GSKDVPNTSL NSTGQPVTKL ATEKDILSTS QTVTLBPFT VEGTSASLND 540
 GSKTVLRSPH NMLSGTAESL NTVSITEYEE EBLTTSFKLD TGAEDSSGSS PATSAIPFIS 600
 ENISQGYIFS SNMPTITTD VLIPEBANA SEDSTSGGSE ESKDPSMKG NVWFPSTDI 660
 60 TAQPDVQSGR ESFLQINYTE IRVDESEKTT KSFSAQVMS QGSPVTDLEM EHYSTPAYFP 720
 TEVTPHAFTP SSRQQLDVST VNVVYSQTTQ FVYNGETPLQ PSYSSEVEFL VTPLLLENQI 780
 LNTTPAASSS DSALHATPVF PSVDVSFESI LSSYDGAPLL PFSSASFSEE LFRHLHTVSQ 840
 ILPQVTSATE SDKVLPHASL PVAGGDLLE PSLAQYSDVL STTHAASETL EPGSESGLVY 900
 KTLMPGQVEP PSSDAMHAR GSGPEPSYAL SDNEGSGQIF TVSYSSAIPV HDSVGVITYQG 960
 65 SLFSGPSHIP IPKSSLLTPT ASLLQFTHAL SGDGENSGAS SDSEFLLPDT DGLTALNISS 1020
 PVSVAEFTYT TSVFGDDNKA LSKSEIYGN ETELQIPSPN EMVYPSSESTV MPMYDNVNR 1080
 LNASLQRTSV SISSTGKMPF GSLARTTKV FDHEISQVPE NNFSVQPTBT VSQASGDTSL 1140
 KPVLSANSEP ASSDPASSEM LSPSTQLLFY ETSASFSTEV ILQPSFQASD VDTLLKTVLP 1200
 AVPEDPILVE TPVKDKISST MLHLIVSNSA SSENMLHSTS VPVFDVSPTS HMHGASLQGL 1260
 70 TISTASEKYE PVLLKSESGH QVPSLYEND ELFQTANLEY NQAHFPRGRH VFATPVLSD 1320
 EPLNLTINKL IHEDEILTST KSSVTGKVEA GIFTVASDTF VSTDHEVPIC NGEVAITVS 1380
 FHRDGSVTST KLLFPKATS ELSSHAKSDA GLVGGGEGDG TDDDDDDDD DRGSDGLSIE 1440
 KMSCCSSYRS SQKVMNDGD THENSLMDQN NPISYSLSEN SEEDNRVTSV SSDSQTGMDR 1500
 75 SPKSPSAGS LSKHNDGKE ENDIQTGSA LPLSPESKAW AVLTSDSSG SGQGTSDSLN 1560
 ENETSTDFSF ADTNEKADG ILAAGDSEIT PGFPQSTSS VTSENSEVFFH VSEASNSNS 1620
 HESRIGLAEG LESEKKAVIP LVIVSALTFI CLVVLVGLI YWRKCPQTAH FYLEDSTSPR 1680
 VISTPPTPIF PISDDVGATP IKEFPKHVAD LHASSGFTTE FELTKBFYQE VQSCIVDLGI 1740
 TADSSNBEIN KHKNRVINTV AYDSRVRKLA QLAEKDGKLT DYINANYVDG YNRPKAYIAA 1800
 80 QGPLKSTAD PKNMIHNNV EVIVMITNLV EKGRRKCDQY WPDGSGEYB NPLVTKQSV 1860
 VLAYXTVBNF TLNRTKLKKG SQKGRPSGRV VTQVHYTQWP DMGVPEYSLP VLTFRKAA 1920
 AKRHAVGPVV VHCBSAGVRT GTYIVLDSML QQIQHEGTVN IFGFLKHIS QRNYLVQTEK 1980
 QYVPHDITLV SALSKEIEV LDSHIAVYN ALLIPGPAGK TKLEKQFQLL SQSNIQQSDY 2040
 SAALKQCNRE KNTTSIIPV ERSRVGISSL SGEGTDYINA SYIMGYQSN EPIITQHELL 2100
 HTIKDFWEMI WDHNAQIVVM IPDGGQMAED EFVYVPMKOE PINCESPKVT LMAEHKCL 2160

NEEKLIQDF ILBATQDDYV LEVRHFQCPK WPNFDPISK TPELISVIKE EAANRDGPMI 2220
 VDEHGGVTA GTFCALITLM HQLEKENSVD VYQVAKMINL MRPGVFADIE QYQFLYKVL 2280
 SLVSTRQEN PSTSLDSNGA ALPDGNIAES LESLV

5

Seq ID NO: 373 DNA sequence
 Nucleic Acid Accession #: built from NM_002851
 Coding sequence: 148-4518

10	1	11	21	31	41	51	
	CACACATACG	CACGCACGAT	CTCATTCTGA	TCTATACACT	GGAGGATTAA	AACAAACAAA	60
	CAAAAAAACC	ATTTCCCTTG	CTCCCCCTCC	CTCTCCACTC	TGAGAAGCAG	AGGAGCCGCA	120
	CGGCGAGGGG	CGGCAGACCG	TCTGGAAATG	CGAATCCTAA	AGCGTTTCCT	CGCTTGCAAT	180
	CAGCTCCTCT	GTGTTTGCCG	CCTGGATTGG	GCTAATGGAT	ACTACAGACA	ACAGAGAAAA	240
15	CTGTGTGAAG	AGATTGGCTG	GTCCTATACA	GGAGCACTGA	ATCAAAAAAA	TTGGGGAAAG	300
	AAATATCCAA	CHGTATATAG	CCCAAAACAA	TCTCCTATCA	ATATTGATGA	AGATCTTACA	360
	CAAGTAAATG	TGAATCTTAA	GAAACTTAAA	TTTCAGGGTT	GGGATAAAAC	ATCATTGGAA	420
	AACACATTCA	TTCAATAACAC	TGGGAAAAACA	GTGGAAATTA	ATCTCACAIA	TGACTACCGT	480
	GTGAGCGGAG	GAGTTTCTGA	AATGGTGTTC	AAAGCAAGCA	AGATAACTTT	TCACTGGGGA	540
20	AAATGCAATA	TGTATCTCTA	TGGATCAGAG	CATAGTTTAG	AAGGACAAAA	ATTTCCACTT	600
	GAGATGCAAA	TCTACTGCTT	TGATGCGGAC	CGATTTTCAA	GTTTTGAGCA	AGCAGTCAAA	660
	GGAAAAGGGA	AGTTAAGAGC	TTTATCCATT	TTGTTTGAGG	TTGGGACAGA	AGAAAATTTG	720
	GATTTCAAGG	CGATTATTTA	TGGAGTCGAA	AGTGTTAGTC	GTTTTGGGAA	CGAGGCTGCT	780
25	TTAGATCCAT	TCATATCTGT	GAACCTTCTG	CCAAACTCAA	CTGACAAGTA	TTACATTTAC	840
	AATGGCTCAT	TGACATCTCC	TCCCTGCACA	GACACAGTTG	ACTGGATTGT	TTTAAAGAT	900
	ACAGTTAGCA	TCTCTGAAGG	CCAGTTGGCT	GTTTTTTGTG	AGTTCCTTAC	AATGCACAAA	960
	TCGTGTTATG	TCAATGCTAT	GGACTACTTA	CAAAACAATT	TTGAGAGACA	ACAGTACAAG	1020
	TTCTCTAGAC	AGGTGTTTTT	CTCATACTCT	GGAAAGGAAG	AGATTTCATG	AGCAGTTTGT	1080
	AGTTCAAGAG	CAGAAAATGT	TCAGGCTGAC	CCAGAGAATT	ATACCAAGCT	TCTTGTATCA	1140
30	TGGGAAAGAC	CTCGAGTCTG	TTATGATACC	ATGATTGAGA	AGTTTGCACT	TTTGTACCAG	1200
	CAGTTGGATG	GAGAGGACCA	AACCAAGCAT	GAATTTTTGA	CAGATGGCTA	TCAAGACTTG	1260
	GGTGCTATTC	TCAATAATTT	GCTACCCAT	ATGAGTTATG	TTCTTCAGAT	AGTAGCCATA	1320
	TGCACTAATG	GCTTATATGG	AAAATACAGC	GACCAACTGA	TTGTGACAT	GCCTACTGAT	1380
	AATCCTGAAC	TTGATCTTTT	CCCTGAATTA	ATTGGAATCG	AAGAAATAAT	CAAGGAGGAG	1440
35	GAAGAGGGAA	AAGACATTGA	AGAAGGCGCT	ATTGTGAATC	CTGGTAGAGA	CAGTGCTACA	1500
	AACCAAAATCA	GGAAAAGGA	ACCCAGATT	TCTACCAACA	CACACTACAA	TGCAATAGGG	1560
	ACGAAATACA	ATGAAGCCAA	GACTAACCGA	TCCCCAACAA	GAGGAAGTGA	ATTCTCTGGA	1620
	AAGGGTGATG	TTCCCAATAC	ATCTTTAAAT	TCCACTTCCC	AACCACTCAC	TAAATTAGCC	1680
	ACAGAAAAGG	ATATTTCCCT	GACTTCTCAG	ACTGTGACTG	AACCTGCCAC	TCAACTGTG	1740
40	GAAGTACTCT	CAGCCTCTTT	AAATGATGGC	TCTAAACTCG	TTCTTAGATC	TCCACATATG	1800
	AACTTGTGGG	GGACTGCGAG	ATCCTTAAAT	ACAGTTTCTA	TAAACAGATA	TGAGGAGGAG	1860
	AGTTTATTGA	CCAGTTTCAR	GCTTGATACT	GGAGCTGAAG	ATTCTTCAGG	CTCCAGTCCC	1920
	GCAACTTCTG	CTATCCCAAT	CATCTCTGAG	AACATATCCC	AAGGATATAT	ATTTTCTCCC	1980
	GAAGACCCAG	AGCAATAAAC	CTTATATGTC	CTTATACCAAG	AATCTGCTAG	AAATGCTTCC	2040
45	GAAGATTCAA	CTTCATCTGG	TTCAAGAGAA	TCACTAAAGG	ATCCTTCTAT	GGAGGGAAT	2100
	GTGTGGTTTC	CTAGCTCTAC	AGACATAACA	GCACAGGCGG	ATGTTGGATC	AGGCAGAGAG	2160
	AGCTTTCCTC	AGACTAATTA	CACCTGAGATA	CGTGTGATG	AATCTGAGAA	GACAAACCAAG	2220
	TCCCTTTCTG	CAGGCCCAAT	GATGTCAACG	GGTCCCTCAG	TTACAGATCT	GGAATAGCCA	2280
	CATTATTTCA	CTTTTGCTCA	CTTCCCAACT	GAGGTAACAC	CTCATGCTTT	TACCCCATCC	2340
50	TCCAGACAAC	AGGATTGGTT	CTCCACGCTC	AACGTGGTAT	ACTCGCAGAC	AACCCAAACG	2400
	GTATACAATG	CAGAGGCCAG	TAAATAGTAG	CATGAGTCTC	GTATTGGTCT	AGCTGAGGGG	2460
	TTGGAATCCG	AGAAGAGGCG	AGTTATACCC	CTTGTGATCG	TGTCAGCCCT	GACTTTTTATC	2520
	TGCTTAGTGG	TTCTTGTGGG	TATCTCTATC	TACTGGAGGA	AATGCTTTCA	GACTGCACTC	2580
	TTTTACTTAG	AGGACAGTAC	ATCCCTTAGA	GTTATATCCA	CACCTCCAAC	ACCTATCTTC	2640
55	CCAAATTTAG	ATGATGTGGG	AGCAATTTCA	ATAAAGCACT	TTCCAAGACA	TGTATGAGAT	2700
	TTACATGCAA	GTATGAGGTT	TACTGAAGAA	TTTGAGACAC	TGAAGAGGTT	TTACCAGGAA	2760
	GTGAGAGGCT	GTACTGTTGA	CTTAGGTATT	ACAGCAGACA	GCTCCAACCA	CCAGACCAAC	2820
	AAGCACAAGA	ATCGATACAT	AAATATCGTT	GCCTATGATC	ATAGCAGGGT	TAGCTAGACA	2880
	CAGCTTGCTG	AAAAGGATGG	CAAACTGACT	GATATATACA	ATGCCAATTA	TGTTGATGCG	2940
60	TACAACAGAC	CAAAAGCTTA	TATGTCTGCC	CAAGGCCAC	TGAATCCAC	AGCTGAGGAT	3000
	TTCTGAGAG	TGATATGGGA	ACATAATGTG	GAAGTTATTG	TCAATGATAAC	AAACCTCGTG	3060
	GAGAAAGGAA	GGAGAAATAG	TGATCAGTAC	TGGCCCTGCG	ATGGGAGTGA	GGAGTACGGG	3120
	AACCTTCTGG	TCACTCAGAA	GAGTGTGCAA	GTGCTTGCCCT	ATTATACTGT	GAGGAATTTT	3180
	ACTCTAAGAA	ACACAAATAT	AAAAAAGGGC	TCCCAAGAAAG	GAAGACCCAG	TGGACGTGTG	3240
65	GTCAACAGAT	ATCACTACAC	CGAGTGGGCT	GACATGGGAG	TACCAGAGTA	CTCCCTGCCA	3300
	GTGCTGACCT	TGTGAGAAA	GGCAGGCTAT	GCCAGGCGCC	ATGCAGTGGG	GCCTGTGTCT	3360
	GTCCACTGCA	GTGCTGGAGT	TGGAGAGACA	GGCACATATA	TTGTGCTAGA	CAGTATGTTG	3420
	CAGCAGATTG	AACACGAGAG	AACGTGTCAAC	ATATTGGGCT	TCTTAAACAA	CATCGGTTCA	3480
	CAAGAAATAT	ATTTGGTACA	AACGTAGGAG	CAATATGTCT	TCATTCAATGA	TACACTGGTT	3540
70	GAGGCCATAC	TTAGTAAAGA	AACGTAGGTT	CTGGACAGTC	ATATTCAATG	CTATGTTAAT	3600
	GCACCTCTCA	TTCTCGGACC	AGCAGGCCAA	ACAAAGCTAG	AGAAACAAAT	CCAGCTCTCT	3660
	AGCCAGTCAA	ATATACAGCA	GAGTGAATAT	TCTGCAGCCC	TAAAGCAATG	CAACAGGGAA	3720
	AGAATCGAA	CTTCTCTAT	CATCCCTGTG	GAAAGATCAA	GGGTGGGCAAT	TTCTATCCCTG	3780
	AGTGGAGAA	CGACAGACTA	CATCAATGCC	TCCATATACA	TGGGCTATTA	CCAGAGCAAT	3840
75	GAATTCATCA	TTACCCAGCA	CCCTCTCCTT	CATACCATCA	AGGATTTCTG	GAGGATGATA	3900
	TGGGACCAT	ATGCCCACT	GGTGGTTATG	ATTCCTGATG	GCCAAACAT	GGCAGAGGAT	3960
	GAATTTGTTT	ATGGGCCAAA	TAAAGATGAG	CCTATAAAT	GTGAGAGCTT	TAAAGTCACT	4020
	CTTATGGCTG	AGAAGACAAA	ATGTCTATCT	AAAGAGGAAA	AACCTTATAT	TCAGGACTTT	4080
	ATCTTAGAAG	CTACACAGCA	TGATTATGTA	CTTGAAGTGA	GGCACTTTCA	GTGCTCTAAA	4140
	TGGCCAAAT	CAGATAGCCC	CATTAGTAAA	ACTTTTGAAC	TTATAAGTGT	TATAAAGGAA	4200
80	GAAGCTGCCA	ATAGGAGTGG	GCCTATGATT	GTTCATGATG	AGCATGGAGG	AGTGAGGCA	4260
	GGAACTTTCT	GTGCTCTGAC	AACCCCTATG	CACCAACTAG	AAAAGAAAA	TTCCGTGGAT	4320
	GTTTACCAAG	TAGCCAGGAT	GATCAATCTG	ATGAGGCCAG	GAGTCTTTGC	TGACATTGAG	4380
	CAGTATCAGT	TTCTCTACAA	AGTGATCTCT	AGCCTTGTGA	GCACAAGGCA	GGAAGAGGAT	4440